ABSTRACT

Title of Document:	BIOLOGY AND EVOLUTION OF CHROMALVEOLATE PROTISTS
	John James Miller, Ph.D. 2010
Directed By:	Professor Charles F. Delwiche, Department of Cell Biology and Molecular Genetics

Dinoflagellates and haptophytes are both prominent members of the marine phytoplankton and are considered chromalveolates. The interactions of the parasitic dinoflagellate *Amoebophrya* with its host dinoflagellate *Akashiwo sanguinea* were studied using cell biological techniques. The free-swimming dinospore stage of *Amoebophrya* has two flagella, trichocysts, striated strips, condensed chromatin resembling heterochromatin, and electron dense bodies. When entering the host cytoplasm and again when entering the host nucleus, the electron dense bodies appear in a tube of microtubules close to the surface of the host or its nucleus. Host entry is inhibited by cytochalasin D implying a role for microfilament polymerization in the entry process. While in the host cytoplasm, *Amoebophrya* appears to be separated from the host cytoplasm by two membranes. After entering the host nucleus, the parasite grows and

undergoes mitosis forming a multinucleated trophont. The mastigocoel is an internal cavity that contains flagella and becomes the outside of the parasite after it leaves the host. This study indicates that the mastigocoel forms as a result of vesicle fusion. Eventually, *Amoebophrya* fills the host nucleus and takes on a beehive appearance. The beehive stage contains numerous trichocysts and striated strips. The level of chromatin condensation in intracellular trophonts is highly variable. It then exits its host as a multinucleated vermiform shaped creature, which then splits up into individual infective dinospores. A phylogenomic pipeline was designed to analyze the genome and evolutionary history of the haptophyte *Emiliania huxleyi*. It appears to have genes linking it to three lineages: heterokonts, green algae, and red algae. Genes with shared phylogenetic affinities appear to fit into limited functional categories and be physically localized in the genome. The phylogenetic affinities of *E. huxleyi* with the green algae may be an artifact of the much greater number of sequenced genomes from the Viridiplantae (=plants+ green algae) when compared to the rhodophytes. The evolutionary history of *E. huxleyi* is still unclear although they do seem to be similar in many ways to heterokonts and are generally believed to have red algae derived plastids.

BIOLOGY AND EVOLUTION OF CHROMALVEOLATE PROTISTS

By

John J. Miller

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2010

Advisory Committee:
Professor Charles F. Delwiche
Professor D. Wayne Coats
Professor Najib El-Sayed
Professor Charles Mitter
Professor David Mosser

Acknowledgments

Thanks to Dr. Charles F. Delwiche for providing a positive environment for me to work, access to the *Emiliania huxleyi* genome, and plentiful useful advice. Thanks to Dr. D. Wayne Coats for getting me started working on *Amoebophrya* and providing me with cultures. Thanks to the rest of my committee including Dr. Najib El-Sayed, Dr. Charles Mitter, and Dr. David Mosser for their advice and constructive comments on my dissertation. Thanks to Dr. Betsy Read for giving me access to the *E. huxleyi* genome data. I would also like to acknowledge people in the Delwiche lab: Dr. Ruth Timme for her advice regarding phylogenetic analysis, Letitia Cosbert for helping maintain cultures, Will MacTurk for proofreading much of my dissertation, Dr. Sara Handy and Gregg Mendez for valuable conversations about science and graduate school and the other members of the lab for helpful comments and encouragement. Thanks to Dr. Cristian Castillo-Davis and Dr. Tsvetan R. Bachvaroff for advice regarding possible pitfalls of my pipeline. Thanks to George A. Miller for proofreading my dissertation. Any remaining errors are my own. Thanks to Tim Maugel and Jan Endlich for their advice regarding my ultrastructural work. Thanks to my family, particularly my wife, Ana, for their patience and support as I undertook this long journey. This work was supported in part by NSF grant #EF0629624 "AToL: an integrated approach to the phylogeny of dinoflagellates. The sequencing of the *E. huxleyi* genome was done at the Joint Genome Institute. The work conducted

by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No.DE-AC02-05CH11231.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1: Introduction

Chapter 2: Ultrastructure of *Amoebophrya* sp. and its changes during the couse of infection

Chapter 3: *Emiliania huxleyi* phylogenomics

Chapter 1: Introduction

The study of protists is important because they make up a major component of life on earth and have large effects on the health of the planet. Eukaryotic life is currently divided into five supergroups, which include the Opisthokonts, Plantae, Excavates, Rhizaria, and Chromalveolates. Opisthokonts include animals and fungi as well as protists. The Plantae include plants and protists. The other three supergroups: the Excavates, Rhizaria, and Chromalveolates are composed entirely of protists (Keeling *et al*. 2005). Most eukaryotic organisms are protists and most of these are unicellular, but they have profound global impacts. For example, there were approximately 247 million cases of malaria, a disease caused by protists in the genus *Plasmodium*, and about one million deaths attributed to this disease in 2006 (WHO 2008). In addition, harmful algal blooms result in reduced biodiversity and toxins accumulating in the food chain that sometimes lead to human death. Harmful species of algae come from several groups of protists including several dinoflagellates as well as a diatom and a haptophyte (Erdner *et al*. 2008). Furthermore, because many protists are phototrophs, they make a large contribution to global primary productivity (Falkowski *et al*. 2004) and are involved in global climate change (Falkowski *et al*. 1998). Some haptophytes known as coccolithophorids, make calcium carbonate scales which result in the sedimentation of carbonate to the bottom of the ocean thus removing carbon dioxide, a greenhouse gas, from the atmosphere. At times haptophytes have produced so many scales that they left behind large geological structures such

as the White Cliffs of Dover and other chalk sediments of the Cretaceous. These are composed largely of coccoliths as well as forams produced by foraminifera, another group of protists (Mitchell *et al*. 1997). Thus, understanding protists is essential to understanding human health and the functioning of the planet. The protists I focused on are within the eukaryotic supergroup Chromalveolata. This supergroup includes the alveolates, haptophytes, heterokonts and cryptomonads (=cryptophytes). These organisms are unified because some members of each group possess secondary plastids containing chlorophylls a and c (Cavalier-Smith 1999).

The possession of secondary plastids is not the only example of symbiosis in this group. Symbiotic relationships in which one of the members is harmed are parasitic relationships (Saffo 1993). Among the alveolates, most apicomplexans (Oborník *et al*. 2009) and some dinoflagellates (Coats 1999) are parasites, as are some heterokonts, for example *Phytophthora* spp. (Hardham 2007). Mutualistic relationships involve cases in which both members benefit (Saffo 1993). Some dinoflagellates, such as *Symbiodinium*, have mutualistic relationships with invertebrates (Yellowlees *et al*. 2008). There is even at least one apicomplexan, *Chromera velia*, which has a mutualistic relationship with a cnidarian host (Moore *et al*. 2008). The third form of symbiosis is commensalism in which one member benefits and the other is not harmed (Saffo 1993).

The origins of plastids

The origins of primary plastids have been traced back to the phagocytic engulfment of a cyanobacterium by a eukaryotic cell (Whatley *et al*. 1979). This was followed not by digestion but the incorporation of the cyanobacterium into the host cell as genes from the endosymbiont were transferred to the host nucleus and the symbiont became a true organelle (Martin *et al*. 1998). The stroma of primary plastids are surrounded by two membranes (Cavalier-Smith 1981). There are three main lineages in possession of a primary plastid: Glaucophyta, Rhodophyta, and Viridiplantae (Delwiche 1999; Deschamps and Moreira 2009; Whatley *et al*. 1979). The plastids of Glaucophytes contain single thylakoids, phycobilisomes, chlorophyll a, and a reduced genome typical of plastids (Delwiche 1999). The Glaucophyte plastid still retains a remnant bacterial cell wall hinting further at its cyanobacterial origin (Delwiche 1999; Whatley *et al*. 1979). Rhodophyte plastids also have single thylakoids, chlorophyll a, and phycobilisomes but lack any remnant of a bacterial cell wall (Cavalier-Smith 1981; Delwiche 1999). Viridiplantae plastids contain chlorophyll a and b, have stacked thylakoids, and lack phycobilisomes.

Secondary plastids arose when a plastidic eukaryotic cell was engulfed by another eukaryote. Secondary plastids usually have a stroma surrounded by four membranes although some have only three. The outer two are usually thought to correspond to the plasma membrane of the engulfed cell and the endosomal membrane of the secondary host (Cavalier-Smith 1999; Delwiche

1999). In cases in which there are only three, the secondary plastid may have been acquired through myzocytosis. Myzocytosis is a process in which a cell punctures another cells plasma membrane with a feeding tube and withdraws the cytoplasm through the tube. Thus a secondary plastid acquired through myzocytosis would lack the membrane corresponding to the plasma membrane of the original primary plastid-containing cell (Delwiche 1999) although some believe that originally there was a fourth membrane and it was lost or that the two outer membranes became one (Cavalier-Smith 1999). Two groups of organisms have secondary plastids with chlorophylls a and b indicating they are derived from the eukaryotic green algae lineage. The chlorarachniophytes have a plastid with four membranes and a residual nucleus, called a nucleomorph, between the outer two and the inner two periplastidal membranes. The phototrophic euglenoids surround their plastid stroma by only three membranes. Euglenoids are known to feed myzocytotically (Cavalier-Smith 1981; Delwiche 1999). Several groups of organisms have plastids containing chlorophylls a and c all of which have either four or three periplastidal membranes. These are thought to be red algal derived secondary plastids. These include haptophytes, heterokonts, cryptomonads, and dinoflagellates (Cavalier-Smith 1981; Delwiche 1999). The rest of this dissertation focuses on organisms with secondary red algal derived plastids and their relatives so I will describe each of these major groups of organisms in some detail.

Cryptomonads

Cryptomonads (=cryptophytes) show the clearest signs of having Rhodophyte derived secondary plastids. Their plastids have a reduced nucleus, called a nucleomorph, in the periplastidal space between the outer and inner pairs of membranes of each plastid (figure 1). The nucleomorph genome has a similar arrangement of genes as that of the genome of the red algal *Porphyra purpurea* (Douglas and Penny 1999) implying common ancestory. The periplastidal space contains eukaryotic ribosomes labeled by a eukaryote specific rRNA probe that also labels the nucleolus of the nucleomorph giving further indication that the periplastidal space is the remnant cytoplasm of a eukaryotic cell (McFadden 1990). Also within the periplastidal space are starch

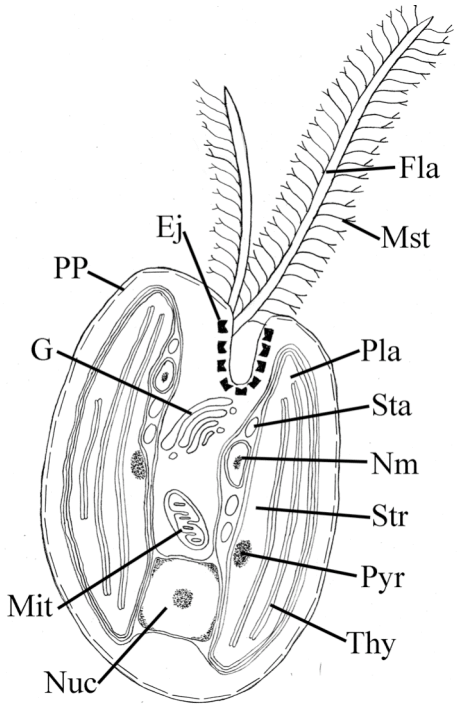


Figure 1: Drawing of a typical member of the genus *Cryptomonas*. Flagellum (Fla); Mastigoneme (Mst); Plastid (Pla); Starch granule (Sta); Nucleomorph (Nm); Stroma (Str); Pyrenoid (Pyr); Thylakoid stack (Thy); Nucleus (Nuc); Mitochondrion (Mit); Golgi apparatus (G); Periplast plate (PP); Ejectisome (Ej). After Santore 1985.

grains, the major form of carbohydrate storage by cryptomonads (figure 1)(Graham and Wilcox 2000). Like red algae and cyanobacteria, cryptomonad plastids contain phycobiliproteins. Unlike cyanobacteria and red algae, which have their phycobiliproteins in phycobilisomes on the stromal surface of their thylakoid membranes (Gantt 1975), cryptomonad phycobiliproteins are found within the thylakoid lumen (Gantt *et al*. 1971) attached to the thylakoid membrane (Ludwig and Gibbs 1989). These protists possess chlorophylls a and c (Cavalier-Smith 1981; Delwiche 1999). Thylakoids are in stacks of two in each plastid stroma along with a pyrenoid (Graham and Wilcox 2000). Cryptomonads are biflagellate cells; the longer flagellum has two rows of bipartite tubular hairs (Adl *et al*. 2005) called mastigonemes and the other flagellum has one row of bipartite mastigonemes (Clay and Kugrens 1999). The flagella emerge near a gullet lined by extrusive organelles called ejectisomes (figure 1), which are structured like rolls of ribbon that telescope out during ejection (Graham and Wilcox 2000). Cryptomonads can be found in both marine and freshwater habitats (Shalchian-Tabrizi *et al*. 2008).

Heterokonts

Heterokonts are a highly morphologically diverse group of organisms. They include macroscopic kelp, unicellular biflagellate chrysophytes, silica frustule bearing diatoms, and parasites such as *Phytophthora* spp. (Graham and Wilcox 2000). The heterokonts are defined primarily by the morphology of their flagella. Heterokont refers to the two unequal flagella of these protists as illustrated in the chrysophyte shown in figure 2. Typically a heterokont cell has a

long immature flagellum with mastigonemes and a shorter mature flagellum that lacks mastigonemes. Heterokonts are sometimes referred to as stramenopiles. A stramenopile refers to the tripartite flagellar hairs on the immature flagella (strameno = straw)(figure 2)(Andersen 2004). Some heterokonts lack flagella over much of their life cycle, but the sperm of many of these organisms have a heterokont configuration of flagella. For example, the sperm of the macroscopic intertidal algae, *Fucus* spp. (Bouck 1969), and the oomycete *Phytophthora parasitica* (Reichle 1969) are biflagellate with one flagellum possessing mastigonemes. Some heterokonts diverge from this pattern, however. Diatoms usually lack flagella, but male gametes of centric diatoms

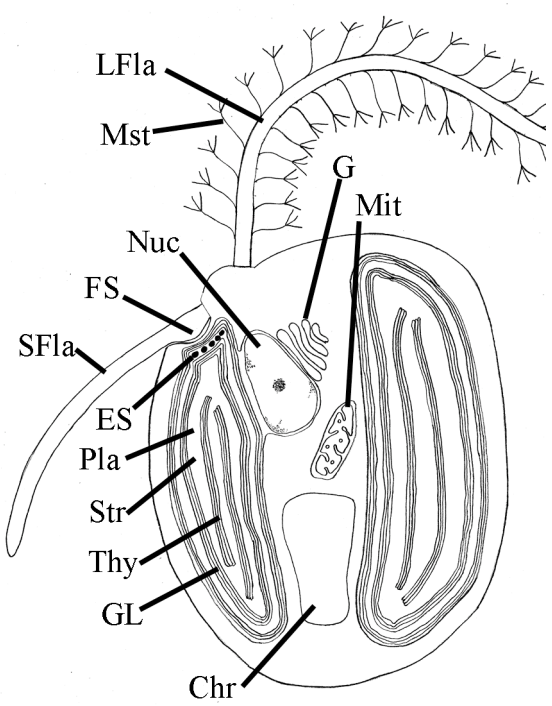


Figure 2: Drawing of a typical chrysophyte heterokont. Long flagellum (LFla); Mastigoneme (Mst); Short flagellum (SFla); Flagellar swelling (FS); Nucleus (Nuc); Eye spot (ES); Plastid (Pla); Stroma (Str); Thylakoid (Thy); Girdle lamella (GL); Chrysolaminaran vesicle (Chr); Golgi apparatus (G); Mitochondrion (Mit). Redrawn from Lee (1989).

have one flagellum with mastigonemes but with an aberrant 9 + 0 configuration of microtubules (Andersen 2004) and cells in meiosis II that will give rise to male gametes also have one flagellum at each spindle pole (Manton *et al*. 1970). Pennate diatoms are not known to ever have flagella. The gametes of the Pinguiophycean, *Glossomastix chrysoplasta*, have one flagellum without mastigonemes (O'Kelly 2002). The shorter flagellum is frequently autofluorescent and likely contains a flavin-like molecule. A flagellar swelling of the shorter flagellum fits into a depression in the cell near a carotenoid containing eyespot usually situated inside a plastid (figure 2). It is thought that the autofluorescence of the flagellum and the eyespot are involved in phototaxis (Jékely 2009). Many heterokonts are heterotrophic, but the plastids of those that have them generally share certain characteristics. Heterokont plastids contain chlorophylls a and c and the accessory pigment fucoxanthin (Strain *et al*. 1944). They have thylakoids in stacks of three and four membranes surround the stroma. A stack of thylakoids, called a girdle lamella, encircles the periphery of the stroma (figure 2)(Andersen 2004; Cavalier-Smith 1981; O'Kelley 2002). Carbohydrates are stored in cytoplasmic vesicles as chrysolaminaran (Graham and Wilcox 2000).

Haptophytes

The Haptophyceae get their name from the presence of a haptonema, a microtubule supported non-flagellar appendage, and are also well known for their production of calcified scales referred to as coccoliths (figure 3)(Andersen 2004;

Manton 1967). The haptonema is used in feeding by some haptophytes such as *Chrysochromulina hirta* (Kawachi *et al*. 1991; Lewin 1992), but other haptophytes, like *Prymnesium patelliferum,* have been observed feeding and apparently do not use their haptonema (Tillmann 1998) so it is possible that some haptonema have other functions such as attachment. Coccolithophorids are haptophytes, which surround themselves with calcified scales called coccoliths. Many non-calcified haptophytes surround themselves with organic scales (Graham and Wilcox 2000)(figure 3). In coccolithophorids, haptonema are frequently reduced as in *Ochosphaera neapolitana* (Frensel and Probert 2005) or absent as in *Emiliania huxleyi* (=*Coccolithus huxleyi*)(Klaveness 1972). A pair of flagella when present are of equal length and do not have mastigonemes (Andersen 2004). Some cells, particularly those with coccoliths,

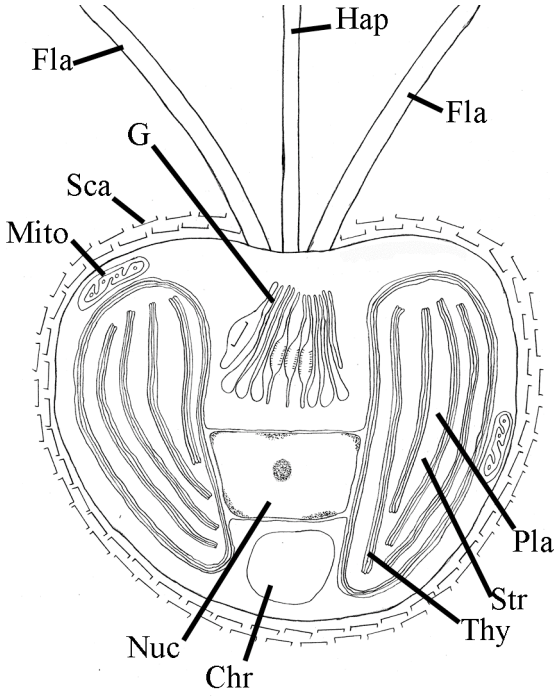


Figure 3: Drawing of a typical haptophyte. Haptonema (Hap); Flagellum (Fla); Golgi apparatus (G); Scales (Sca); Mitochondrion (Mito); Nucleus (Nuc); Chrysolaminaran vesicle (Chr); Thylakoid (Thy); Stroma (Str); Plastid (Pla). Redrawn from Lee 1989).

do not have flagella although other members of the same species may have flagella. In the case of *O. neapolitana*, some coccolith bearing cells have flagella, but in the case of *E. huxleyi* the flagellate cells lack coccoliths instead having organic scales (Green *et al*. 1996). Haptophyte plastids are similar in many ways to those of heterokonts in having four bounding membranes (figure 3), chlorophylls a and c, and having fucoxanthin. In addition, the fucoxanthin derivatives 19' -hexanoyloxyfucoxanthin and 19'-butanoyloxyfucoxanthin are considered characteristic of haptophytes although some heterokonts also have these pigments. Haptophyte plastids differ from those of heterokonts in that they lack a girdle lamella (Andersen 2004; Cavalier-Smith 1981) but like heterokonts, haptophytes store carbohydrates as chrysolaminaran in cytoplasmic vesicles (figure 3)(Graham and Wilcox 2000).

Apicomplexans

The Apicomplexa include organisms like *Plasmodium*, the pathogen that causes malaria. Apicomplexans are named for their apical complex set of organelles including the micronemes, rhoptries, and dense granules (figure 4). The organelles of the apical complex are used by these parasites to gain entry to host cells (Baum *et al*. 2008; Dubremetz *et al*. 1998; Preiser *et al*. 2000). Micronemes are released first and are involved in adhesion to host cells. Rhoptries are released second, contain hydrophobic material, and are involved in formation of the parasitoforous vacuole membrane. Dense granules are then exocytosed and modify the parasitoforous vacuole space (Baum *et al*. 2008; Dubremetz *et al*. 1998; Preiser *et al*. 2000). Host cell entry is

powered by a parasite actin and myosin driven mechanism (Baum *et al*. 2008; Sibley 2004). Despite their parasitic life style, at least some apicomplexans have plastids (figure 4) referred to as apicoplasts. The apicoplast is generally thought to be surrounded by four membranes and has descended from a red algae. While it has no photosynthetic function, this organelle is important to the parasite for its function in the synthesis of fatty acids, isoprenoids, iron-sulfur proteins, and heme. It is considered a good drug target, because animals lack plastids and use different mechanisms for the acquisition and synthesis of the biomolecules made in the apicoplast (Lim and McFadden 2010).

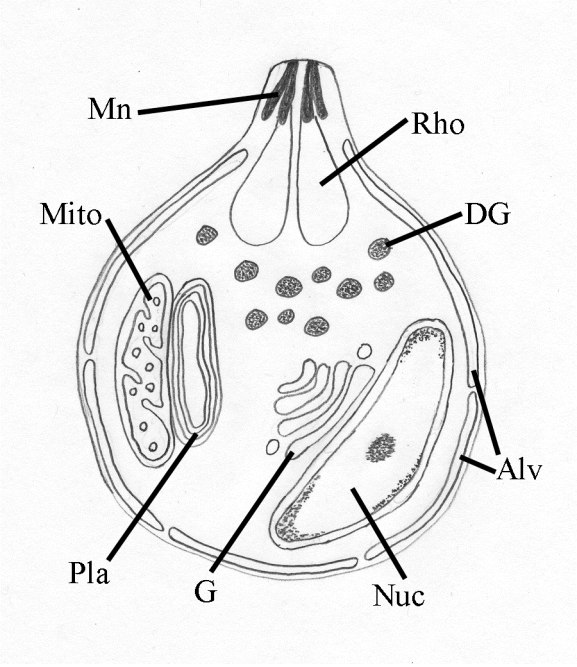


Figure 4: Drawing of *Plasmodium* sp. Microneme (Mn); Rhoptry (Rho); Dense granule (DG); Mitochondrion (Mito); Plastid (Pla); Golgi apparatus (G); Nucleus (Nuc); Alveoli (Alv). After Preiser *et al*. 2000 modifications from Struck *et al*. 2005.

Dinoflagellates

Dinoflagellates are best known as a source of toxins such as paralytic shellfish poison (PSP)(Wang 2008; Wong and Kwok 2005). Like their alveolate relatives the apicomplexans and ciliates, the dinoflagellate cortex has a set of

flattened vesicles referred to as alveoli (figure 5). Some dinoflagellates have cellulosic plates within their alveoli. The cortex, including alveoli, is sometimes called the amphiesma. The dinoflagellate nucleus has unique features. Although eukaryotic they lack normal histones and have permanently condensed chromosomes (figure 5). During mitosis, they send spindle microtubules through tunnels lined by nuclear envelope and attach the microtubules to chromosomes through an intact nuclear envelope (Bhaud *et al*. 2000). These biflagellate organisms have a transverse flagellum encircling the cell in a groove called a girdle or cingulum and another longitudinal flagellum starting in a groove called the sulcus (figure 5)(Graham and Wilcox 2000). The transverse flagellum has a striated strand (= paraxonemal

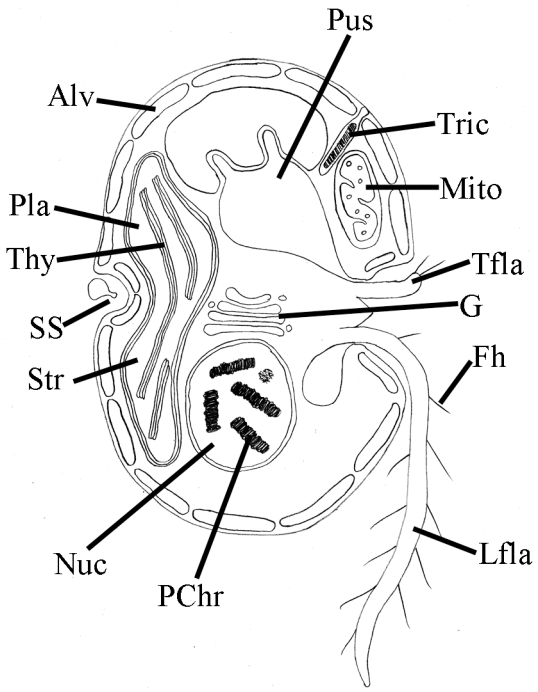


Figure 5: Drawing of a dinoflagellate. Alveoli (Alv); Plastid (Pla); Thylakoid (Thy); Stroma (Str); Striated strand (SS); Nucleus (Nuc); Permanently condensed chromosomes (PChr); Longitudinal flagellum (Lfla); Flagellar hair (Fh); Transvers flagellum (Tfla); Golgi apparatus (G); Mitochondrion (Mito); Trichocyst (Tric). Redrawn from Taylor 1980.

rod) associated with the axoneme and the longitudinal flagellum has a paraxial rod associated with its axoneme with both flagella having flagellar hairs (Dodge and Greuet 1987; Leadbeater and Dodge 1967; Taylor 1980). Dinoflagellates as a group seem to have gotten plastids from several different sources. Most dinoflagellate plastids contain chlorophylls a and c, and contain the unique xanthophyll, peridinin (Haidak *et al*. 1966; Iglesias-Prieto *et al*. 1993; Strain *et al*. 1944). One explanation for the three membranes rather than four is that the plastid could have been acquired through myzocytosis a process by which the plasma membrane of a prey cell is punctured and the contents ingested without the plasma membrane (Delwiche 1999; Schnepf 1985). Other dinoflagellates, including *Karenia brevis*, *Karenia mikimotoi*, and *Karlodinium micrum*, have plastids containing chlorophylls a and c with the carotenoids 19'-butanoyloxyfucoxanthin, 19'-hexanoyloxyfucoxanthin, and fucoxanthin, a pigment composition characteristic of haptophytes (Bjornland *et al*. 2003; Schnepf and Elbrachter 1999). Furthermore, a few dinoflagellates such as *Glenodinium foliaceum* and *Peridinium balticum* have the remnants of a diatom that retains its own nucleus. These dinoflagellates have chlorophylls a and c with the carotenoid, fucoxanthin, typical of heterokonts (Schnepf and Elbrachter 1999).

Amoebophrya

Amoebophrya is a genus of parasitic flagellates that can limit the growth of dinoflagellate populations including some harmful species (Coats *et al*. 1996; Coats and Bockstahler 1994; Kim *et al*. 2004; Nishitani *et al*. 1985). In some cases *Amoebophrya* may be a more important control than grazing by

microzooplankton (Montagnes *et al*. 2008). Depending on the strain, this parasite can be limited to one host species (Coats *et al*. 1996; Coats and Park 2002) or may be able to infect multiple hosts (Kim 2006). Phylogenetic studies place *Amoebophrya* between the dinoflagellates and the apicomplexans as a group II marine alveolate (Harada *et al*. 2007; Skovgaard *et al*. 2005) implying that it may have characteristics of both apicomplexans and dinoflagellates. *Amoebophrya* spp. have a complex life-cycle in which an infective dinospore invades its host and undergoes multiple rounds of mitosis without cytokinesis eventually forming a syncytium resembling a beehive (Figure 6). The beehive stage has an internal cavity called a mastigocoel in which the flagella grow. Cachon and Cachon (1987) suggest that the mastigocoel forms when a sheet of cytoplasm encases the central core of parasite cytoplasm. While exiting the host the parasite inverts itself so the membrane lining the mastigocoel becomes the plasma membrane of the multinucleated vermiform shaped parasite. The vermiform separates into

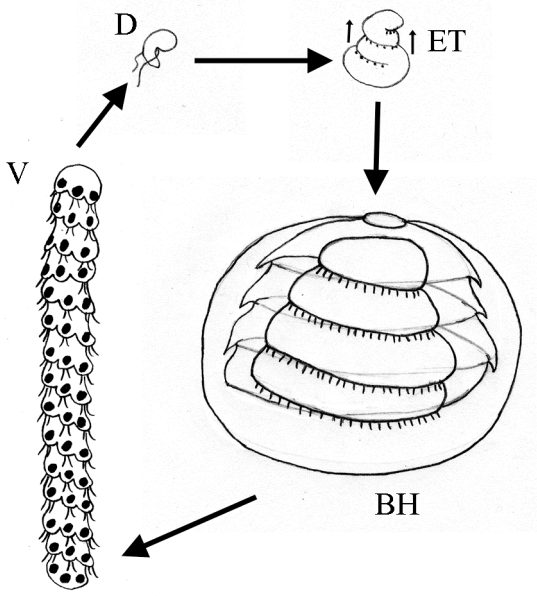


Figure 6: *Amoebophrya* life-cycle. Dinospore (D); Early trophont (ET); Beehive (BH); Vermiform (V). After Cachon 1987.

multiple infective dinospores within minutes of exiting the host. This entire process takes two to three days (Cachon and Cachon 1987; Coats and Bockstahler 1994; Kim *et al*. 2004).

Chromalveolate biology and phylogenetics

The protists I have focused on are in the chromalveolate supergroup. Haptophytes, cryptophytes, heterokonts, and alveolates (include ciliates, dinoflagellates, and apicomplexans) are considered chromalveolates because of their possession of secondary red algal derived plastids. Cavalier-Smith (1999) proposed that because it should be hard to acquire a plastid there was only one such acquisition and the organisms in possession of a secondary red algal derived plastid form a monophyletic group. This became known as the "Chromalveolate hypothesis." The common ancestory of the plastids has received strong support (Le Corguillé *et al*. 2009). The relationship of the host organisms has been more controversial. Some chromalveolates, particularly basal ones lack plastids and there have been several separate acquisitions of plastids within the dinoflagellates making plastid gain seem like less of a rare event (Bodyl 2005). Furthermore, molecular phylogenies support the relationship of heterokonts with alveolates and the relationship of cryptophytes and haptophytes, but for the chromalveolates to form a monophyletic clade, the rhizaria (a group of amoeboid protists including foraminifera and heliozoans) must be included in the chromalveolates (Burki *et al*. 2007). More recently green algal derived genes were found in diatom genomes and interpreted as meaning

that there was a separate green algal endosymbiosis prior to the red algal endosymbiosis or that diatoms actually descended from a green alga (Moustafa *et al*. 2009).

I am using two approaches to understand the biology of the chromalveolates. First I am using ultrastructural techniques to examine the parasitic dinoflagellate *Amoebophrya* sp. Second, I am using phylogenomic methods to study the coccolith producing haptophyte *Emiliania huxleyi*. *Amoebophrya* sp. parasitizes another dinoflagellate, in this case a phototroph, *Akashiwo sanguinea*. *Amoebophrya* falls phylogenetically between the apicomplexans and the dinoflagellates (Harada *et al*. 2007; Skovgaard *et al*. 2005). My work has allowed me to test several hypotheses about *Amoebophrya*. One is that *Amoebophrya* has permanently condensed chromosomes. A second hypothesis is *Amoebophrya* has plastids. Another is that *Amoebophrya* shares common mechanisms of infection with apicomplexans. I used transmission electron microscopy (TEM) and scanning electron microscopy (SEM) to study the life-cycle of *Amoebophrya* sp. I collected TEM and SEM data on infective dinospores separated from their host and TEM data on infected hosts at 10 min, 12 hr, 24 hr, 36 hr, and 48 hr time points. I also used the cytoskeletal inhibitors cytochalasin D (inhibits microfilament polymerization) and nocodazol (inhibits microtubule polymerization) to answer questions about how *Amoebophrya* enters its host. My work has indicated that *Amoebophrya* does not have permanently condensed chromosomes. No evidence of plastids was found. Although *Amoebophrya* does not seem to have an apical complex, it does seem to used

microfilaments as it enters its host and it has electron dense structures that appear to be involved in infection and could be homologous to structures of the apical complex of apicomplexans.

The second approach I am taking is a phylogenomic approach with *Emiliania huxleyi* as a target organism. Given that *E. huxleyi* is likely to have had gene inputs from at least two organisms: a red algae and the ancestral host organism with a possible third a green algae, I designed an analytical pipeline that keeps track of separate phylogenetic signals within the genome of *E. huxleyi*. The pipeline finds homologs using a pairwise reciprocal best hit (prbh) blast (Altschul *et al*. 1997) strategy then looks for clusters of orthologous groups (COGs). A COG is a group of sequences in which each sequence has a prbh relationship with at least three other sequences in the group (Tatusov *et al*. 1997). Each COG was then aligned using MUSCLE (Edgar 2004) and each alignment subjected to phylogenetic analysis using RAxML (Stamatakis 2006). The pipline found approximately 1200 phylogenetic trees. The bipartitions were counted and the trees were sorted based on which bipartition included *E. huxleyi*. Several concatenated alignments were made based on the sorted trees. Phylogenetic analyses based on separate concatenated alignments supported *Emiliania* grouping with the heterokonts, the green algae, and the red algae. Follow up analyses indicated that the proteins grouped by evolutionary history have related functions. The proteins that grouped *E. huxleyi* with heterokonts were enriched in plastid related genes while the proteins that grouped *Emiliania* with greens were enriched in informational proteins. Both the heterokont lineage

genes and the green lineage genes are distributed on the scaffolds differently than would be expected based on the lengths of the scaffolds implying that genes with similar evolutionary histories may show some synteny.

Chapter 2: Ultrastructure of *Amoebophrya* sp. and its changes during the course of infection.

Abstract

Amoebophrya spp. are syndinian dinoflagellates that parasitize protists including other dinoflagellates and can limit host populations. Prior ultrastructural investigations have focused primarily on the intracellular stages of the life cycle, and relatively little is known about the early stages including attachment to and penetration of the host cell. I undertook a study of these early stages, making use of both ultrastructural and inhibitor studies to elucidate the early stages of the *Amoebophrya* life cycle. The *Amoebophrya* dinospore has a bulbous episome and a narrow hyposome. Two flagella emerge at the base of the episome. Each of which has a thin accessory structure, but only one has flagellar hairs. The dinospore nucleus has a thick layer of heterochromatin around its periphery. Striated plates and dinoflagellate-type trichocysts are found in vesicles around the periphery of the cytoplasm beneath the amphiesma. Electron dense bodies, which sometimes appear coiled, are prominent in the dinospore cytoplasm and appear to be important in infection of its host *Akashiwo sanguinea*. Cytochalasin D was effective in inhibiting host cell entry by *Amoebophrya* whereas nocodazol was not, indicating microfilament but not microtubule polymerization is important during infection. At 12 and 24 h, cytoplasm appears larger and the nucleolus is prominent. At 36 h, multiple nuclei are present and the mastigocoel is beginning to form. At 48 h, the trophont fills most of the host cell and possesses a large

mastigocoel, many flagella, elongated mitochondria and abundant, and abundant trichocysts and striated plates. These results indicate that *Amoebophrya* has important similarities both to more typical dinoflagellates and to apicomplexans.

Introduction

The Alveolates are a relatively well-defined group of protists, comprised of ciliates, apicomplexans, and dinoflagellates (Fast *et al*. 2002; Gajadhar *et al*. 1991; Wolters 1991). Despite their morphological diversity, the alveolates typically possess cortical vesicles ("alveoli") subtending their plasma membrane (Patterson 1999). Ciliates are thought to be the earliest diverging of the alveolates (Fast *et al*. 2002; Gajadhar *et al*. 1991; Wolters 1991). They are known for their complex ciliature and dimorphic nuclei and are important consumers of bacteria and protists in many aquatic ecosystems (Hausmann and Hülsmann 1996). The apicomplexans and dinoflagellates together form a clade that is sister to the ciliates (Fast *et al*. 2002; Gajadhar *et al*. 1991; Wolters 1991). Apicomplexans were, until recently, thought to be entirely parasitic. However, *Colpodella* is a free swimming myzocytotic feeding apicomplexan (Brugerolle 2002; Kuvardina *et al*. 2002; Leander *et al*. 2003; Myl'nikov 2000; Simpson and Patterson 1996) and *Chromera* is a phototroph, which can be cultured without its host coral and is considered most closely related to apicomplexans (Moore *et al*. 2008). Apicomplexans are characterized ultrastructurally by an apical complex specialized for entry into host cells, which is composed of three membranous organelles: micronemes, rhoptries, and dense granules. Micronemes are involved in attachment to a host cell and gliding motility; rhoptries are used in the formation of a parasitophorous vacuole membrane; dense granules then modify the vacuole membrane and contents (Dubremetz *et al*. 1998). Apicomplexans include species that cause several important diseases of humans, livestock, and

other animals, including malaria, toxoplasmosis, and Texas fever (Hausmann and Hülsmann 1996). Dinoflagellates are best known for their role as primary producers and for their dominant role in causing harmful algal blooms (Wong and Kwok 2005). Less well known is that many dinoflagellates are heterotrophic and that many plastidic dinoflagellates are mixotrophic. Furthermore, dinoflagellates can be either ectoparasites or endoparasites (Schnepf and Elbrächter 1992). *Piscinoodinium pillulare* is an ectoparasite of fish and has well-developed plastids making it a mixotrophic parasite (Lom and Schubert 1983). Most dinoflagellates also have a number of unique cytological characteristics, including permanently condensed chromosomes, a lack of normal histones, a distinctive pattern of nuclear division termed “dinomitosis” (Bhaud *et al*. 2000), as well as a striated strand (= paraxonemal rod) associated with their transverse flagellum and a paraxial rod associated with their longitudinal flagellum (Dodge and Greuet 1987).

The dinoflagellates, apicomplexans, and ciliates each constitute a distinctive and diverse clade within the alveolates. However, there are several known organisms that have both apicomplexan- and dinoflagellate-like morphological characteristics. *Perkinsus marinus* (= *Dermocystidium marinum*) was initially discovered as a result of an extensive search for the cause of large-scale death of oysters (Mackin, *et al*. 1950; Villalba *et al*. 2004). The free-swimming stage of *Perkinsus marinus* is an inconspicuous flagellate with rhoptries, micronemes, and a conoid reminiscent of structures found in the apical complex of apicomplexans (Perkins 1976; Villalba *et al*. 2004). At one time

Perkinsus was considered an apicomplexan because of its apical complex (Perkins 1976), but molecular phylogenetic data place it at the base of dinoflagellates (Leander and Keeling 2004; Saldarriaga *et al*. 2003). *Parvilucifera* is another perkinsid that has a potentially large ecological impact because it is an intracellular parasite of dinoflagellates (Noren *et al*. 1999). *Rastrimonas subtilis* (= *Cryptophagus subtilis*), a consumer of cryptophyte algae, is also considered a perkinsid (Brugerolle 2002; Brugerolle 2003). *Colpodella* spp. also possess flagella and an apical complex, but are predators of other protists and are thought to be closer to the apicomplexans based on both ultrastructural and molecular data (Brugerolle 2002; Kuvardina *et al*. 2002; Leander *et al*. 2003; Myl'nikov 2000; Simpson and Patterson 1996). The presence of dinoflagellates with apical complexes and flagellate myzocytotic feeders among the apicomplexans has lead to the proposal that the common ancestor of the dinoflagellates and apicomplexans was a biflagellate with an apical complex (Leander and Keeling 2003; Leander *et al.* 2003). Furthermore, the apicomplexan-dinoflagellate clade has been named the Myzozoa (Cavalier-Smith and Chao 2004). Recently, *Chromera velia*, a phototrophic organism more closely related to apicomplexans was discovered (Moore *et al*. 2008).

With the advent of molecular ecology, it was discovered that there is a large diversity of organisms that fit into two groups which branch early in the dinoflagellate lineage (Groisillier *et al*. 2006). Because they were initially uncharacterized morphologically, the lineages were arbitrarily named 'group I' and 'group II' alveolates. The group I alveolate sequences remained of unknown

morphological affinity for some time, but it has recently been shown that *Dubosquella*, an organism that parasitizes tintinnid ciliates (Coats 1988) as well as other ciliates and dinoflagellates (Cachon 1964) and *Ichthyodinium* an endoparasite of fish eggs, fall within the group I alveolates (Harada *et al*. 2007; Skovgaard *et al*. 2009). The group II alveolates appear to be roughly synonymous with the Syndinids and include the genera *Syndinium*, *Hematodinium*, and *Amoebophrya* (Cachon and Cachon 1987; Skovgaard *et al*. 2005).

The Syndiniales are parasites believed to have diverged early from the dinoflagellate lineage (Cachon and Cachon 1987; Skovgaard *et al*. 2005). *Syndinium* is a parasite of radiolarians (Ris and Kubai, 1974) as well as copepod and amphipod crustaceans (Manier *et al*. 1971; Skovgaard *et al*. 2005; Soyer 1974). These crustaceans are of economic importance because they are a food source for many fish (Day *et al*. 1989). Another genus, *Hematodinium*, is a parasite of decapod crustaceans., Many decapods are commercially important as a human food source (Stentiford and Shields 2005). Lastly, *Amobeophrya* parasitizes a variety of protists, including other dinoflagellates (Cachon and Cachon, 1987). *Amoebophrya* is of particular interest because it infects bloom-forming dinoflagellates, some of which are toxic. *Amoebophrya* has been shown to have a significant impact on some dinoflagellate populations (Coats *et al*., 1996; Coats and Bockstahler 1994; Kim *et al*. 2004; Maranda, 2001; Taylor 1968) sometimes greater than microzooplankton (Montagnes *et al*. 2008). The genus has a complex life cycle that includes both an intracellular and

extracellular phase. After finding and entering a host, the parasite starts to grow either within the host cytoplasm (Coats and Park, 2002) or the host nucleus (Coats and Bockstahler 1994) depending on the host and parasite strain; the *Amoebophrya* trophont undergoes repeated rounds of mitosis, eventually (roughly 2-3 days after infection in cultured material) forming a syncytium that almost completely fills the host. The syncytium then inverts itself as it exits the host to form a short-lived, multinucleate vermiform, which in turn separates into many individual infective dinospores (Cachon and Cachon 1987).

There has been some significant ultrastructural work done on the development of group II alveolates. The development of *Syndinium* spp. within both copepod and radiolarian hosts with emphasis on cell division was studied in detail in the 1970's (Manier 1971; Ris and Kubai 1974; Soyer 1974). The life-cycle of *Hematodinium* sp. within in vitro cultures has been carefully documented (Appleton and Vickerman 1998). Several important studies of the later stages of *Amoebophrya* spp. within its hosts are on record (Cachon and Cachon 1969; Cachon and Cachon, 1970; Fritz and Nass 1992; Maranda, 2001). However, I am unfamiliar with any detailed study of the morphology of free-swimming *Amoebophrya* dinospores or early events in the infection process using TEM.

One group of early diverging endoparasitic dinoflagellates, the perkinsids, has been shown to have apical complex-like organelles (Brugerolle 2002; Noren *et al*. 1999; Perkins 1976). The presence of similar structures in *Amoebophrya* would be of interest, both because it would clarify the relationship between apicomplexans and dinoflagellates, and because it would help elucidate the

evolution of the apical complex organelles used to such devastating effect by organisms such as *Plasmodium* spp. The availability of cultures of *Amoebophrya ceratii* with the host *Akashiwo sanguinea* made it possible to collect ultrastructural data describing the *Amoebophrya* dinospore and early events in the infection cycle such as the penetration of the host pellicle. In addition, the use of cultured material has allowed us to better define time points at which ultrastructural characteristics appear.

Methods

Organism handling

Akashiwo sanguinea from Chesapeake Bay were grown in 15 f/2-Si with soil extract at 20^{o}C, with cool-white fluorescent illumination 75 μmol photons/m^{2}/s on a 15:9 light:dark cycle. Aliquots of *Amoebophrya* culture were transferred into fresh host three times a week. *Amoebophrya* dinospores were purified from whole cultures by gravity filtration using 47 mm Whatman Nuclepore track-etched polycarbonate membranes with 8 μm pores (Coats and Park 2002) (GE Healthcare).

Binding location

Amoebophrya dinospores were incubated with *A. sanguinea* for 5 minutes at a time. Cells were fixed with formalin buffered with $CaCO_3$ at a final formalin concentration of 2% and visualized via autofluorescence (Coats and Bockstahler 1994). *Amoebophrya* dinospores bound to regions of the surface of *A. sanguinea* (episome, hyposome, girdle, sulcus) were enumerated from microscope slides with coverslips (#1.5) using a Zeiss Axioskop with a band pass 450-490 nm excitation filter, 510 nm dichroic mirror, and a long pass 520 nm emission filter at 400X magnification (Carl Zeiss, Inc., NY, USA). Areas of *A. sanguinea* cells were estimated from 3 replicates of 30 cells each photographed with a Zeiss axiocam. Areas of the episome, hyposome, girdle, and sulcus were measured using Zeiss axiovision. Expected numbers of parasites bound to each

major region of the host surface were calculated based on the surface area of each region and statistical significance was tested with a chi-square done in R-2.10.1 (Hornik 2009).

Inhibitor Studies

Cytochalasin D (CytD) was dissolved in DMSO at a concentration of 1mg/ml. The cytD and DMSO were added to cells so the final concentration of cytD was 1.0 μm and of DMSO 0.001% (v/v). The final concentration of *Amoebophrya* dinospores was 10,000/ml and *A. sanguinea* 1000/ml. Three replicates were performed, each of which consisted of three treatments: 1) host pretreated with cytD and parasite with DMSO, 2) host pretreated with DMSO and parasite with cytD, and 3) both host and parasite pretreated with DMSO. Pretreatments lasted 40 min and host and parasite were incubated together for 10 min. Formalin fixation was as above. Parasites were visualized by autofluorescence, and counted on an inverted fluorescence microscope (Leitz, Diavert) with a 450-490 nm excitation and 520 nm emission filters at 400X. Those associated with each of 30 hosts were counted for each replicate of each treatment. Parasites attached to the outside and those fully internalized were tallied separately for each host. Chi-square analysis was performed on the data using R-2.10.1 (Hornik 2009).

Nocodazole was dissolved in DMSO at a concentration of 5 mg/ml. The final concentrations of nocodazole and DMSO after addition to the cells were 1.0 μm and 0.0001% (v/v) respectively. The final concentration of *Amoebophrya* dinospores was 20,000/ml and *A. sanguinea* 2000/ml. Treatments were the same

as for the cytD experiments except that nocodazole was used instead of cytD. Cells were fixed in formalin as described above and autofluorescent parasites were counted on a Zeiss Axioskop at 400X with a 419-461 nm band pass excitation filter, 455 dichroic mirror, and 480 long pass emission filter (Omega Optical, VT, USA). For every replicate of each treatment the parasites associated with 20 hosts were enumerated and results analysed as for the CytD experiments.

Electron Microscopy

For all scanning electron microscopy (SEM) fixes, a 2% OsO_4 15 artificial seawater (Instant Ocean, Aquarium Systems, OH) solution was mixed 1:1 with 15 f/2 containing the organisms. While in fixative, cells were pulled down on 13 mm diameter Whatman Nuclepore track-etched polycarbonate membranes (GE Healthcare) with 1 μm pores within Swinney filter holders (VWR, NJ). While in the Swinney holders, the cells were rinsed in progressively less saline water (10, 5, 0) and dehydrated with an ethanol series (12%, 25%, 50%, 75%, 100%, 100%, 100%). Changes of rinse and ethanol were made at approximately 10 min intervals. Ethanol was removed using a Denton DCP-1 critical point dryer. Samples were then coated with approximately 25 nm of Au:Pd in a DV-503 Denton Vacuum. Specimens were examined with an Amray 1820D SEM and digital images were collected with SEMTech (SEMTech Solutions Inc., MA, USA) and EDAX Genesis (EDAX, NJ, USA) software and processed using Adobe Photoshop, Adobe Photoshop Elements, and Adobe Illustrator (Adobe, San Jose, CA, USA).

Amoebophrya dinospores were presented with *A. sanguinea* for 10 min, 12 h, 24 h, 36 h and 48 h or were kept separate prior to fixation for transmission electron microscopy (TEM). For fixation after 10 min of exposure, *Amoebophrya* was mixed with *A. sanguinea* at a 50:1 parasite: host ratio with a final *Amoebophrya* concentration of 110,000 dinospores/ml. For all other fixations in which hosts were incubated with parasite, the host: parasite ratio was 10:1 with a final concentration of 10,000 *Amoebophrya* dinospores/ml. TEM Fixations of *Amoebophrya* dinospores and *Amoebophrya* incubated with *A. sanguinea* for 10 min were fixed by pulling 1 part 8% glutaraldehyde 0.27 M Sodium Cacodylate 0.027 M MgSO4 solution into a 5 ml syringe with 2 parts 4% OsO4 and injecting it into 5 parts dinoflagellate culture. Fixations were started at room temperature for the first 10 -15 min, but were cooled to 4^{o}C for the last 15-20 min. Cells were rinsed in 0.1 M Sodium Cacodylate and 0.01 M MgSO4 three times over 30 min. The first rinse was started at 4^{o}C but was allowed to warm to room temperature. Host-parasite co-incubations of 12, 24, 36, and 48 h were fixed in the same manner, with the exception that the 8% glutaraldehyde 0.27 M Sodium Cacodylate 0.027 M MgSO4 solution also contained 0.068 M NaCl. These protists were rinsed as above but the rinse solution contained 0.05 M NaCl in addition to 0.1 M Sodium Cacodylate and 0.01 M MgSO4. Cells were dehydrated in ethanol 25-50-75-95-100-100 over 1 h, transitioned through three changes of propylene oxide over 30 min before embedding in Eponate 12 (Ted Pella, Redding CA). In the case of dinospores and 10 min host-parasite incubations, cells were placed in 50% Eponate 12 for 30 min and were then

allowed to sink into a 100% Eponate 12 placed in the bottom of the tubes overnight. Three more changes of 100% Eponate 12 were done the following day before polymerizing at 60^{o}C for 2 days. In the case of time points 12 h or longer, cells went through an Eponate 12 series of 25%-50%-75% at one hour intervals then overnight in 75%. Three changes of 100% were done the following day at two-hour intervals. Silver to gold sections were cut on an AO ultracut microtome with glass knives or a Diatome diamond knife (Diatome, Hatfield, PA, USA). Sections were stained with 0.2% lead citrate (Venable and Coggeshall 1965) followed by 2% uranyl acetate. Sections were viewed on a Zeiss EM10 CA with an 80 kV beam. Images were recorded on 8 cm x 10 cm negative plates then printed onto ILFORD MGIV Multigrade IV RC De Luxe photographic paper (ILFORD PHOTO HARMAN technology Ltd., Mobberley, UK). Prints were then digitized with a CanoScan 8800F Scanner (Canon USA, Inc.) and processed in Adobe Photoshop, Adobe Photoshop elements, and Adobe Illustrator (Adobe, San Jose, CA, USA).

Morphometrics

Areas of trichocysts, electron dense bodies (not lipid bodies), and the *Amoebophrya* cytoplasm as a whole were quantified from TEMs using ImageJ. Morphometric measurements were made from scanned negatives on two groups of 39 cells each: dinospores and parasites that had been fully internalized for ten minutes or less. The Shapiro-Wilk test for normality of distribution and the Wilcoxon-Mann-Whitney hypothesis test were done with R-2.10.1 (Hornik 2009).

Results

Structure of the Dinospore

The free swimming *Amoebophrya* dinospore has a large bulbous episome and a narrow hyposome, which varies in the extent of coiling (1A and B). Each of the two flagella emerges from the base of the episome with the initial directions being posterior and transverse. The transverse flagellum has flagellar hairs (1A, B, and C) and seems to contain an accessory structure that may be a rudimentary striated strand (fig 2A, and B). The longitudinal flagellum is narrower at its distal end (1C) and is considerably shorter than the transverse flagellum. A thin strand is present in the longitudinal flagellum consistent with a paraxial rod

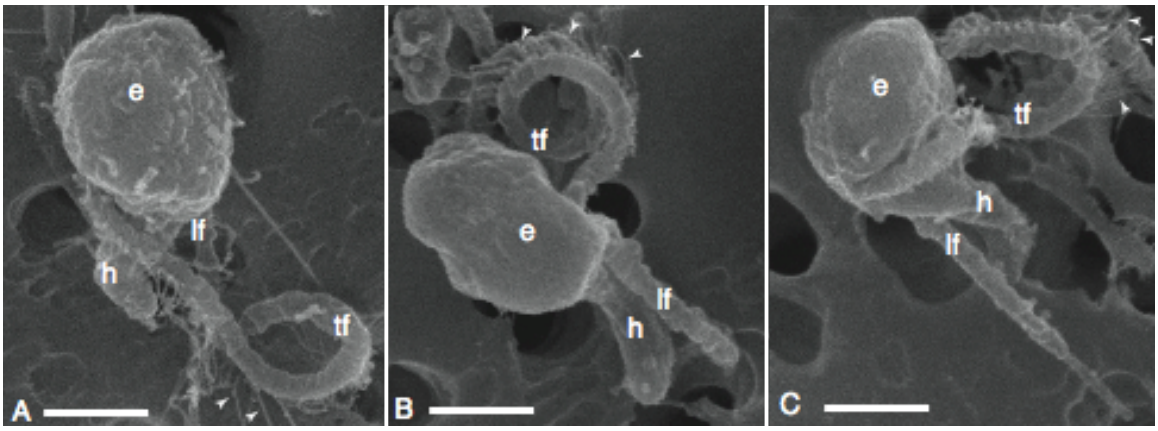


Figure 1. Scanning electron microscope (SEM) images of *Amoebophrya* dinospores viewed dorsally (1A), from the right (1B), and from the left (1C) showing the episome (e), hyposome (h), longitudinal flagellum (lf), transverse flagellum (tf), and flagellar hairs (arrowhead) (bars = 1 μm).

(fig 3A). The angle between the basal bodies is approximately 70^{o} (3C). Longitudinal cortical microtubules are situated in the hyposome (fig 2A, C, and 3A). Trichocysts are found in both the episome and hyposome, but appear to be more numerous in the hyposome (fig 3A and B). Alveoli are flat (fig 2A, 3A, 4A and B) and are distinct from large vesicles that are frequently present directly under the inner amphiesmal membrane. These contain striated strips that appear to be flattened, as opposed to filaments, based on numerous sections through the structures (fig 2A, 3B, D, 4A, B, C, and D). The membrane of these vesicles appears in places to fold in on itself or to be a docking site for the striated structures (fig 2C and 4A). Located primarily in the ventral anterior portion of the episome (fig 2C and 4D) are electron dense structures (fig 4A and C) with a concentric layered internal structure (fig 4D). The nucleus is in the posterior portion of the episome, dorsal to the basal bodies (2A, C, and 4D). A layer of condensed chromatin not separated into distinct chromosomes coats the periphery of the nucleus with little if any condensed chromatin appearing in the central region of the nucleus (fig 2A, C, 4A, B, C, and D). Nucleoli are present in the nucleus in most stages of the life cycle, but could not be identified in dinospores. Filamentous material is found in a pocket in the nuclear envelope (fig 4A). The Golgi apparatus is located anterior-ventrally to the nucleus (2A). A mitochondrion with tubular cristae is located near the basal bodies stretched across the episome hyposome boundary (fig 2A, 2C, 4A, B, and D). A large lipid droplet and sometimes several smaller ones are located anterior to the nucleus

in the dorsal portion of the cell (fig 2A, 4A, B, and D).

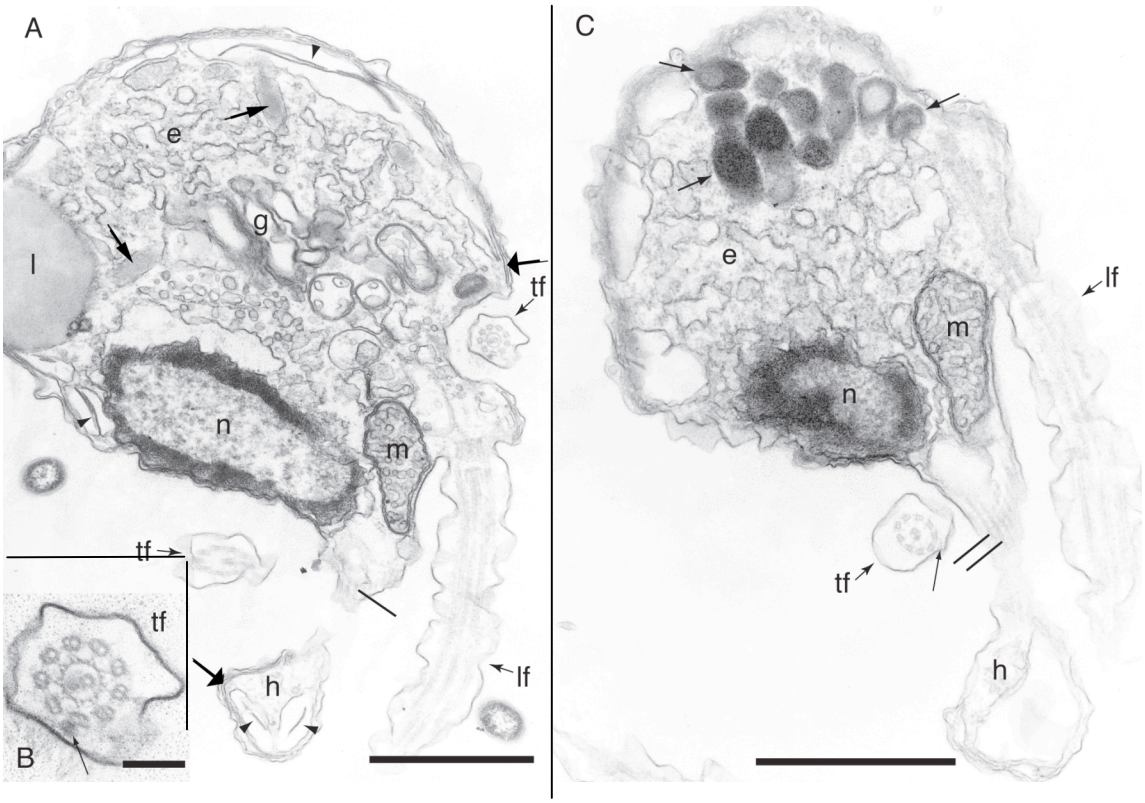


Figure 2. Transmission electron microscope (TEM) images of *Amoebophrya* dinospores. A) Longitudinal section showing episome (e), hyposome (h), longitudinal microtubules of hyposome (line), lipid droplet (l), nucleus (n), mitochondrion (m), Golgi (g), trichocysts (big arrows long heads), transverse flagellum (tf), longitudinal flagellum (lf), alveoli (big arrows wide heads), and striated plates (arrowhead)(bar = 1 μm). B) Basal disk region of transverse flagellum (same as A) showing flagellar hairs and auxillary structure (small arrow long arrowhead) (bar = 0.1 μm). C) Another

longitudinal section showing the episome (e), hyposome (h), longitudinal microtubules of hyposome (line), longitudinal flagellum (lf) with basal body, transverse flagellum (tf) with accessory structure (small arrow long arrowhead) and electron dense bodies (arrows)(1 μm).

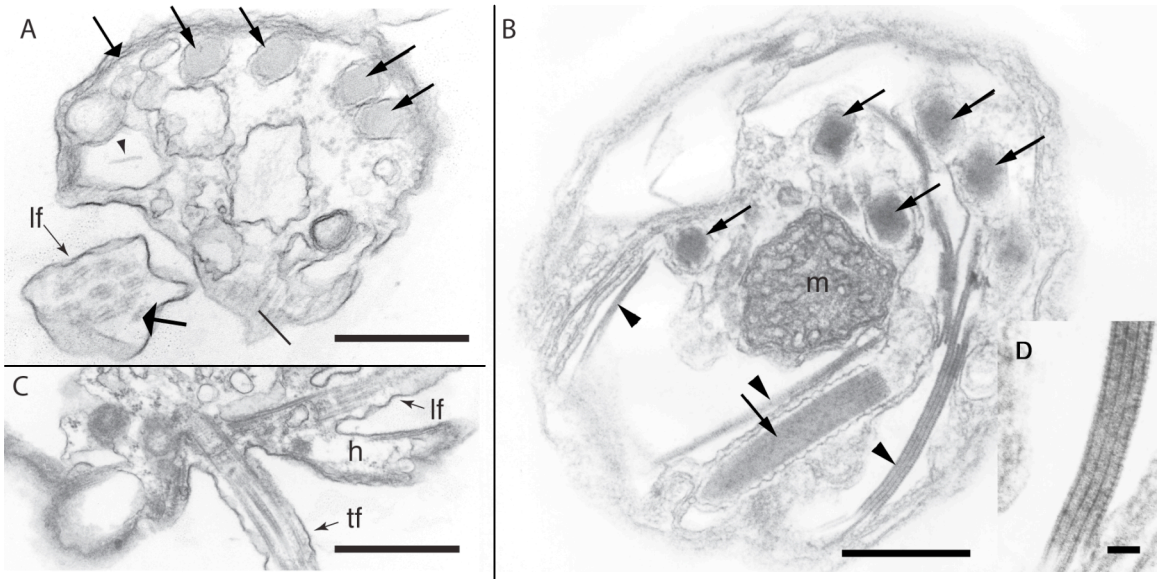


Figure 3. A) Transverse section through hyposome showing trichocysts (thin arrows), striated plates (arrowheads), microtubules (line), alveoli (arrows medium head), and longitudinal flagellum (lf) with accessory structure (arrow with thick head)(bar 0.25 μm). B) Section through dinospore showing trichocysts (small arrows), striated plates (arrowheads), and a mitochondrion (m) (bar = 0.2 μm). C) Shows angle between basal bodies of longitudinal (lf) and transverse flagella (tf)(bar = 0.5 μm). D) Enlargement of stack of striated strips from C (bar = 0.2 μm).

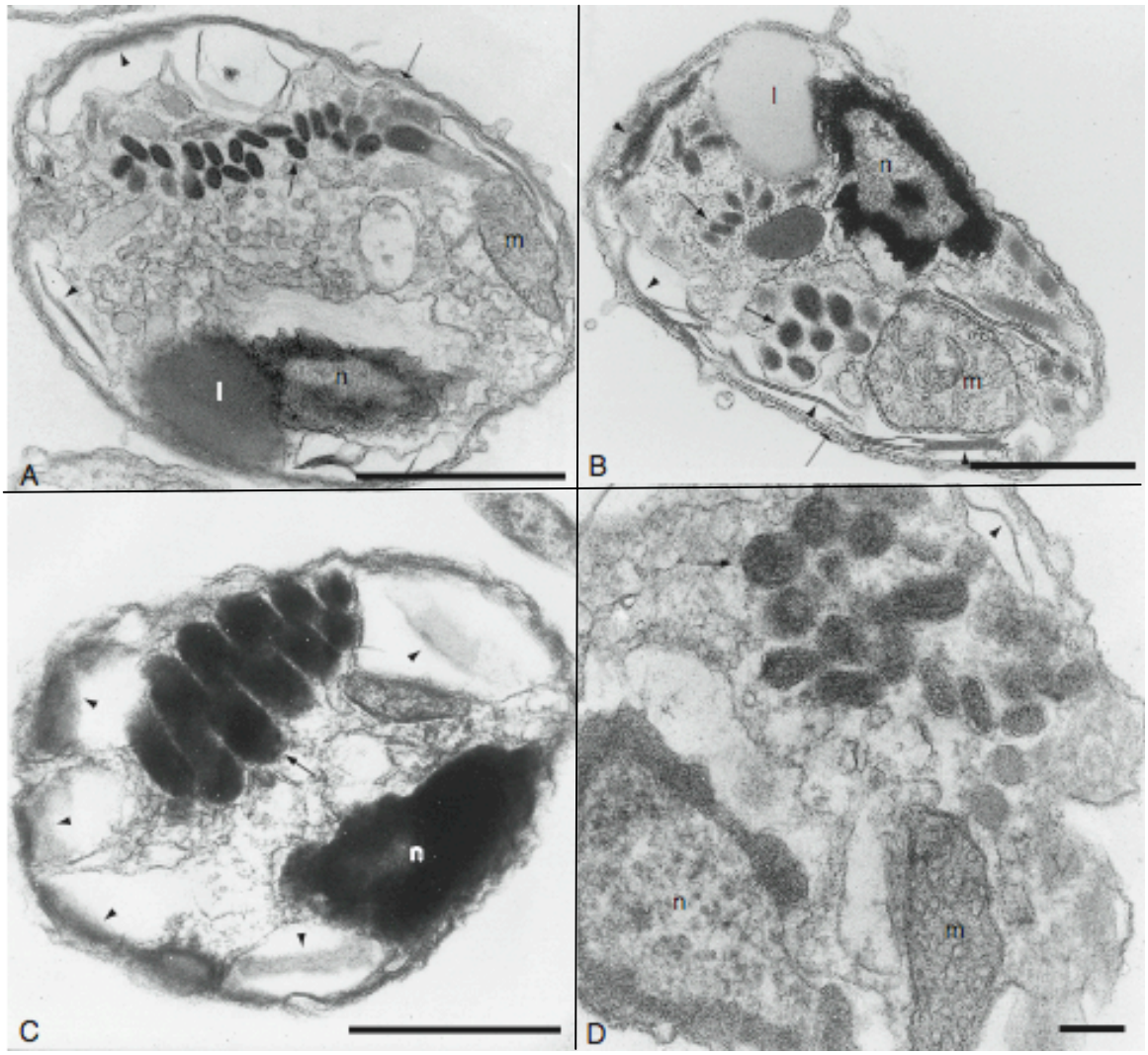


Figure 4. Transmission electron microscope (TEM) images of *Amoebophrya* dinospores. A. Nucleus (n), mitochondrion (m), lipid droplet (l), striated plates (arrowheads), alveoli (arrows wide heads), vesicle infoldings (small arrow) and electron dense bodies (arrows) (bar = 0.5 μm). B. Dinospore showing nucleus (n), mitochondrion (m), lipid droplet (l), striated plates (arrowheads), alveoli (arrows wide heads), and electron dense bodies (arrows) (bar = 0.5 μm). C. Shows spiraling of electron dense body (arrow), planar nature of striated plates (arrowheads), and nucleus (n) (bar = 0.5

μm). D. Shows layered structure of electron dense bodies (arrow), nucleus (n), and mitochondrion (m) (bar = 0.1 μm).

Dynamics of infection

T_0: Although *Amoebophrya* binds to all sections of the surface of *Akashiwo sanguinea*, it is found more frequently in the girdle and sulcus while on the surface of its host (χ^2 = 86.0, df = 3, p < 2.2x10^{-16}) (fig 5 and 6). When *Amoebophrya* initially adheres to the surface of *A. sanguinia*, it retains its flagella, and its flagellar hairs appear to be closely associated with the host surface (fig 7A). Fibers somewhat thicker than the flagellar hairs sometimes emerge from the parasite and attach to the host surface (fig 7A and B). The area taken up by trichocysts is greatly reduced in parasites who had just entered *Akashiwo sanguinea* compared to dinospores, implying they may be used in attachment or discarded during entry (Wilcoxon-Mann-Whitney: U = 1056, p = 2.2 x10^{-05})(fig 8). The parasites lose their flagella and take on a slug-shaped appearance on the surface of their host prior to entering the *A. sanguinea* cytoplasm (fig 7B). At this time, the tip at one end of the parasite can be seen closely adhered to the host surface. TEM images show longitudinal microtubules around the parasite periphery forming a tube near the host amphiesma (fig 9A, B, and C). However, pretreatment of either parasite (χ^2 = 2.2, df = 3, p = 0.53) or host (χ^2 = 0.36, df = 3, p = 0.94) with nocodazole had no significant effect in either case (fig 10A)

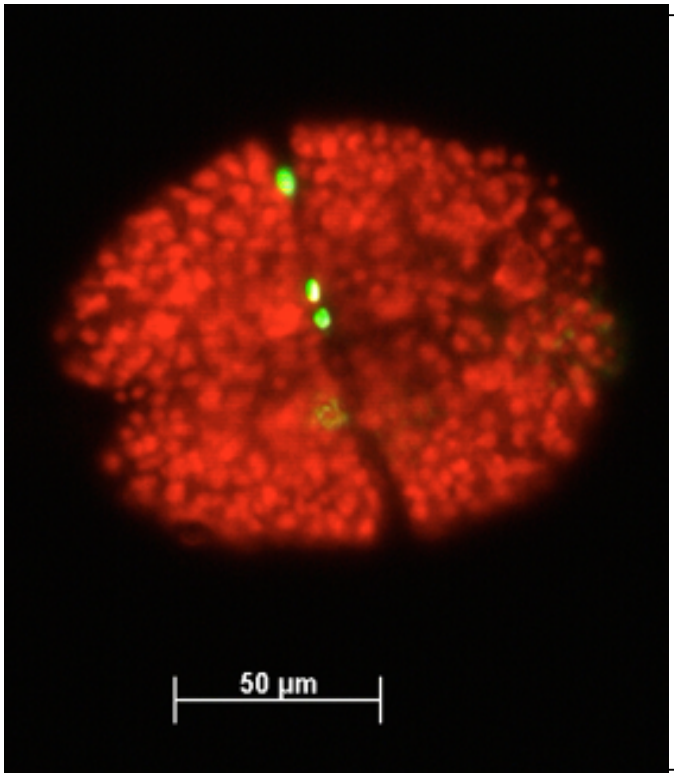


Figure 5. Fluorescence image of Akashiwo sanguinea with three *Amoebophrya* bound to surface within girdle. The parasite appears green. The red is chlorophyll from the host plastids.

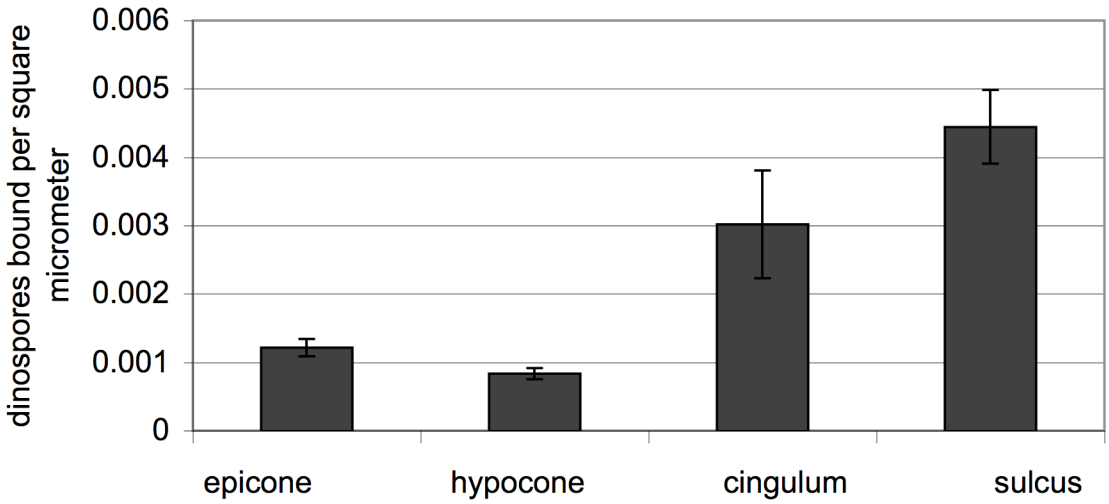


Figure 6 shows *Amoebophrya* bound to major areas of *A. sanguinea* surface per square micrometer. Parasite dinospores preferentially bind to cingulum and sulcus ($\chi^2 = 86.0$, df = 3, $p < 2.2 \times 10^{-16}$).

implying that microtubule polymerization is not essential for host entry. Treatment of *Amoebophrya* dinospores with cytochalasin D reduced the proportion of host associated parasites entering the *A. sanguinea* cytoplasm relative to those parasites bound to the outside of the host (χ^2 = 119.1, df = 3, p < $2.2x10^{-16}$) while pretreatment of host did not make a significant difference (χ^2 = 3.0, df = 3, p = 0.39) implying a role of parasite microfilament polymerization in entry (fig 10B). While fewer parasites reached the host in the cytochalasin D treatments (parasite pre-treatment = 173; host pre-treatment = 155) compared to the DMSO control (=508), implying that cytochalasin D reduced the ability of the parasite to reach its host, clearly the parasites that were stuck to the host amphiesma and those fully internalized were able to reach their host. Therefore comparing ratios of internalized parasites to total host associated parasites among treatments seemed like a valid way to control for the reduced host finding ability. Electron dense bodies are located within the microtubule lined tube (fig 9A, B, and C). The area occupied by electron dense bodies is reduced in early-internalized parasites compared to dinospores, but the difference was not statistically significant (Wilcoxon-Mann-Whitney: U = 882.5, p = 0.11)(fig 11). Membranous material, which could represent either new membrane or material invaginated from the host, is also present within the tube near the host-parasite boundary (fig 9B). The formation of new membrane prior to entering the host is also apparent in the double membrane bound vesicles present in a sac

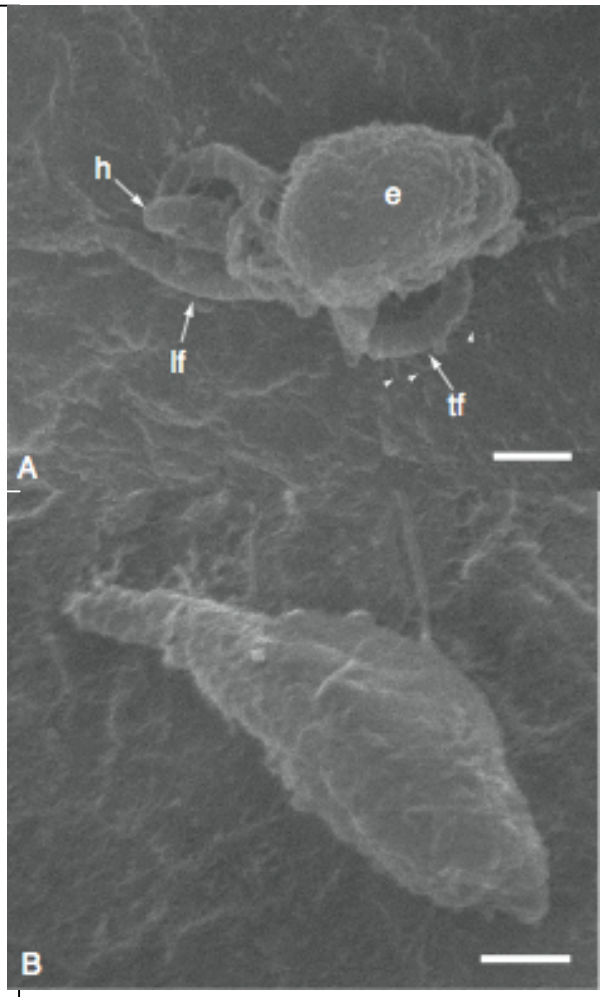


Figure 7. Scanning electron microscope (SEM) images of *Amoebophrya* on the surface of *Akashiwo sanguinea*. A. *Amoebophrya* on the surface of *A. sanguinea* while it still retains its dinospore shape with an episome (e), hyposome (h), longitudinal flagellum (lf), and transverse flagellum (tf) with flagellar hairs (arrowheads) (bar = 1 μm) B. *Amoebophrya* after having taken on a slug shaped appearance (bar = 1 μm).

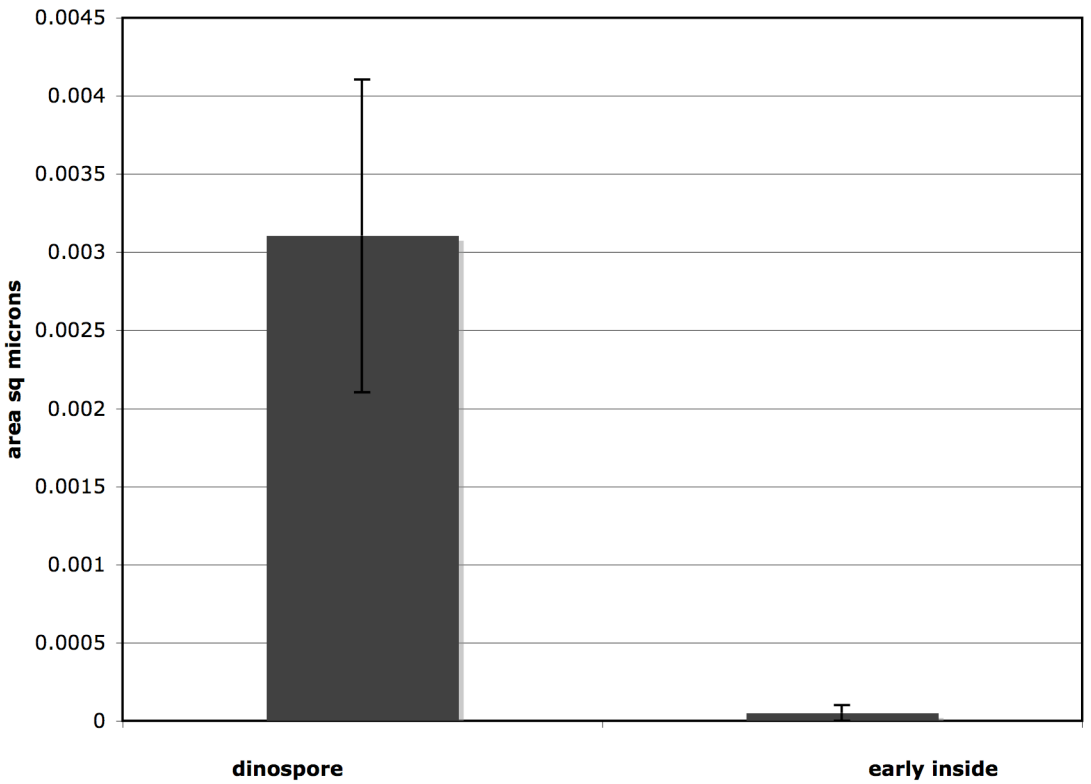


Figure 8. Shows difference in area of trichocysts in *Amoebophrya* dinospores and parasites that had been internalized for less than 10 min Wilcoxon-Mann-Whitney: U = 1056, p = 2.2 $x10^{-05}$.

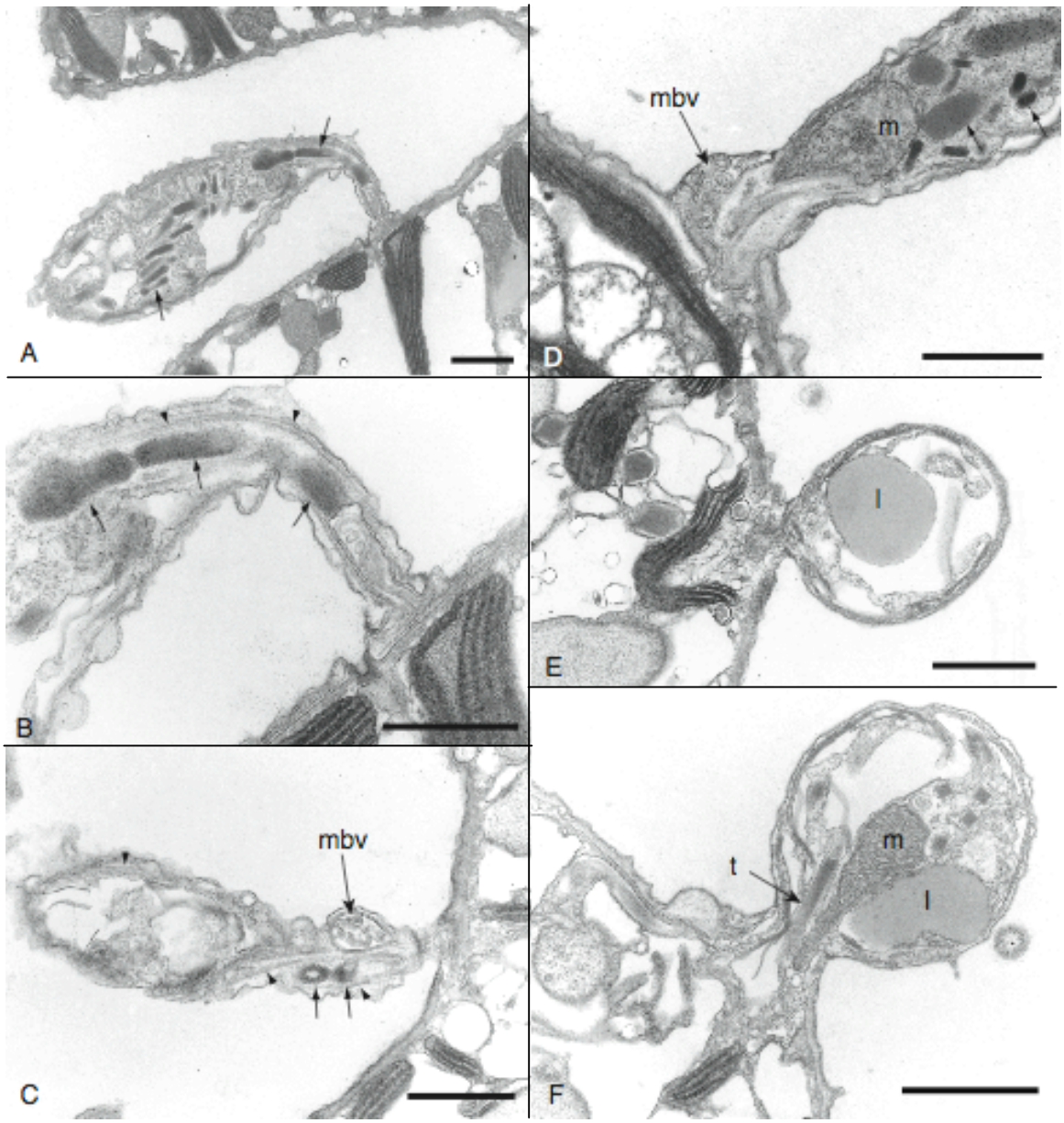
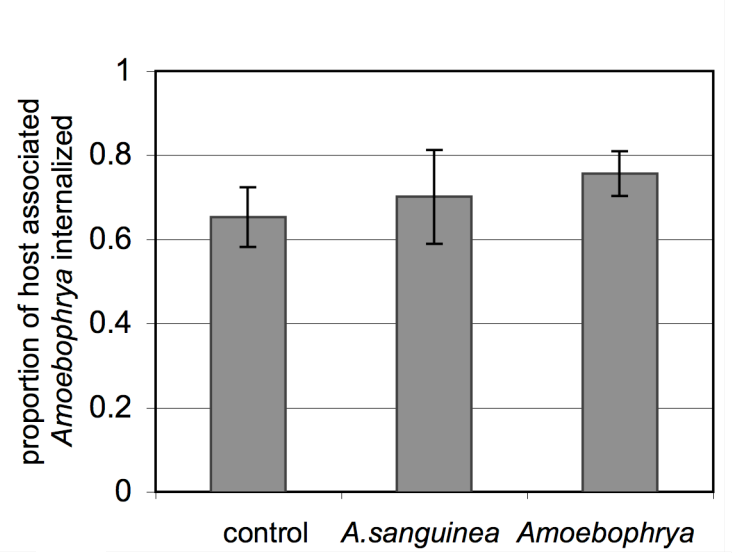


Figure 9. Transmission electron microscope (TEM) images of *Amoebophrya* attached to the surface of *Akashiwo sanguinea*. A. *Amoebophrya* bound in girdle of *A. sanguinea* showing electron dense bodies (small arrow)(bar = 1 μm). B. Enlargement of A. showing the microtubules (arrowheads) associated with electron dense bodies (bar = 1

μm). C. Host associated *Amoebophrya* with multimembrane bound vesicles (mbv) and electron dense bodies associated with microtubules (bar = 1 μm). D. Another *Amoebophrya* showing multimembrane bound vesicles and a mitochondrion (m) as it sticks to *A. sanguinea* (bar = 1 μm). E. *Amoebophrya* with lipid droplet (l) initially inserting vesicles as it enters *A. sanguinea* (bar = 1 μm). F. *Amoebophrya* moving organelles such as a trichocyst (t) and mitochondrion into the cytoplasm of its host (bar = 1 μm).

inside the parasite and outside the parasite near the point of contact with the host (fig 9C and D). Early in the entry process, *Amoebophrya* appears to insert a large number of vesicles into the *A. sanguinea* cytoplasm (fig 9E) prior to inserting its own organelles (fig 9F). During the process of entering its host, the separation of host and parasite cytoplasm seems poorly defined (fig 9F). After penetration is complete, two membranes surround the parasite (fig 12A): a plasma membrane and a parasitophorous vacuole membrane. A mitochondrion is prominent at one end of the parasite (fig 12A and B) and a large vacuole is sometimes seen at the other (fig 12B). Many small vesicles are located near the mitochondrial end of the parasite within the parasite cytoplasm (12A). As seen in the dinospore, condensed chromatin still coats the periphery of the nucleus. Internalized parasites were significantly smaller than dinospores (Wilcoxon-Mann-Whitney: $U = 1120.5$, $p = 1.6 \times 10^{-04}$)(fig 13). A structure that appears to be a tube composed of microtubules containing electron dense bodies is

sometimes seen at one end of the parasite (fig 12C and D), particularly in parasites near the host nucleus.



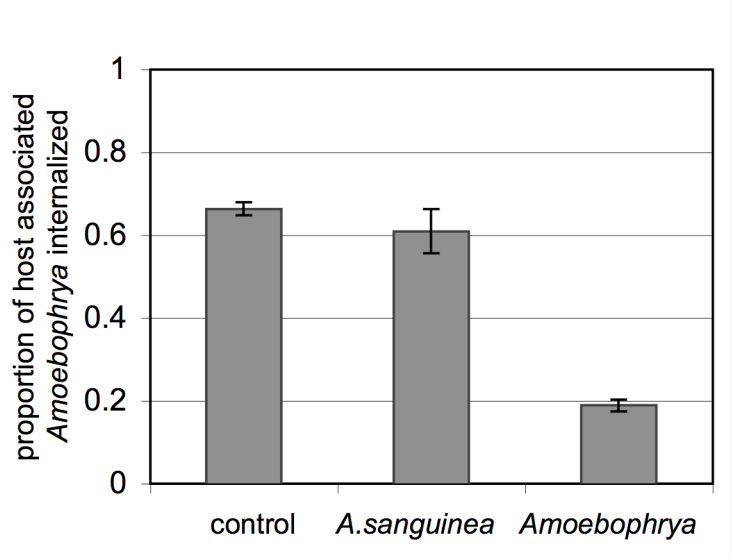


Figure 10. Effects on penetration of host pellicle by *Amoebophrya* of A) Nocodazole pretreatment of *A. sanguinea* (χ^2 = 0.36, df = 3, p = 0.94) or *Amoebophrya* (χ^2 = 2.2, df = 3, p-value = 0.53) and B) Cytochalasin D pretreatment of *A. sanguinea* (χ^2 = 3.0,

df = 3, p = 0.39) or *Amoebophrya* (χ^2 = 119.1, df = 3, p < 2.2x10^{-16}). Control was DMSO in both experiments. Expressed as proportion of internalized parasites per host associated parasite (parasites stuck to the outside and those fully internalized).

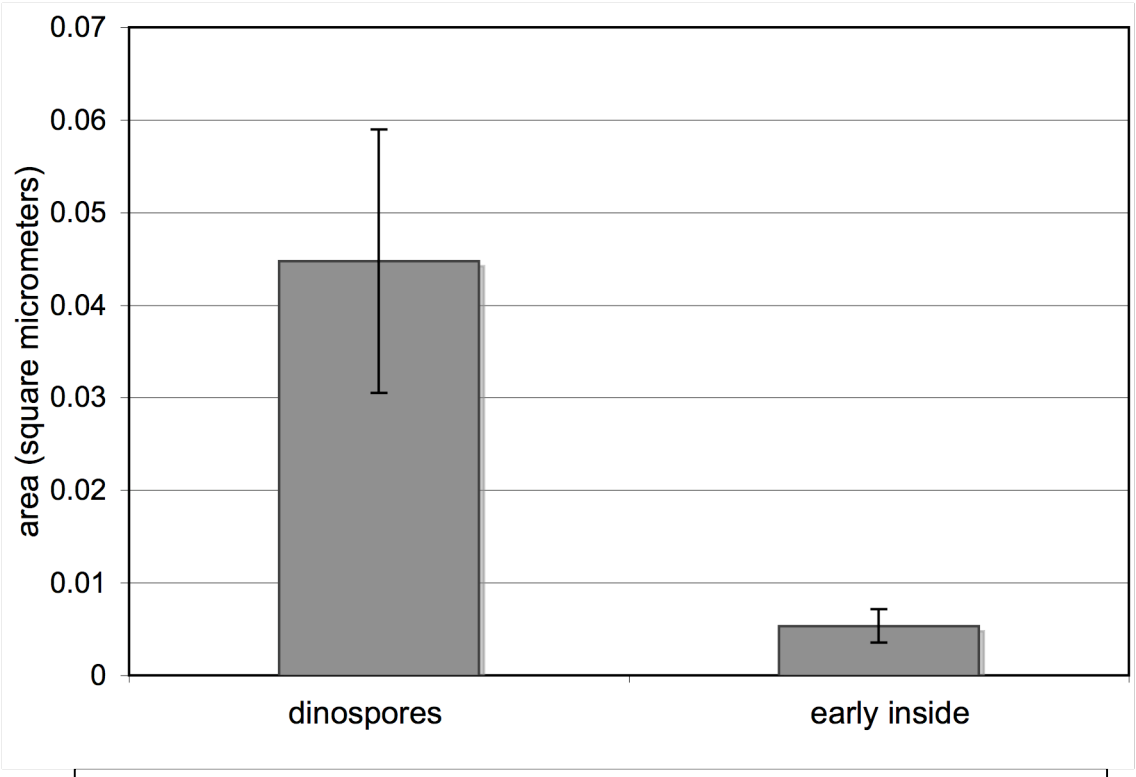


Figure 11. Shows difference in area of electron dense bodies in *Amoebophrya* dinospores and parasites that had been internalized for less than 10 min (Wilcoxon-Mann-Whitney: U = 882.5, p = 0.11).

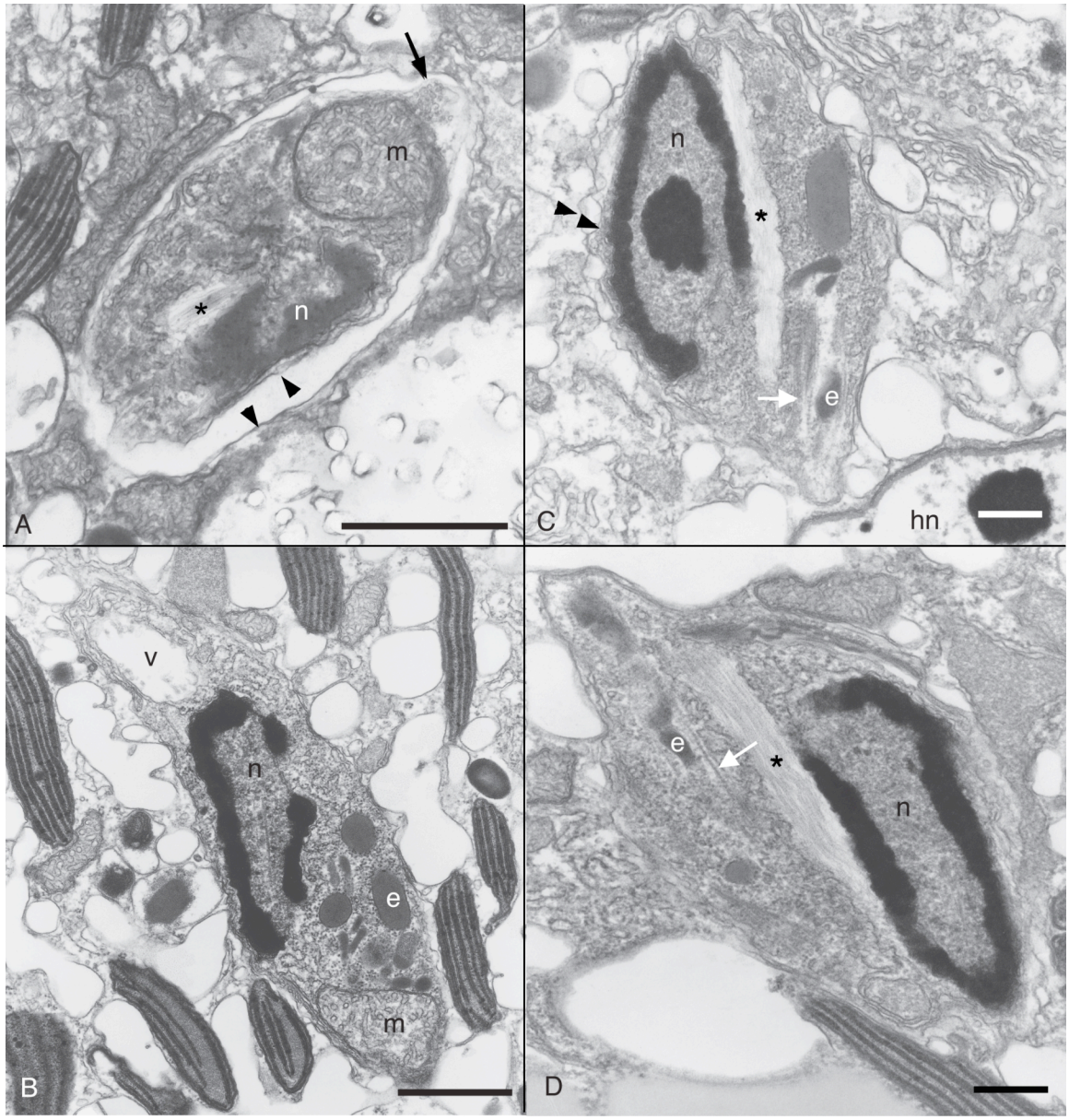


Figure 12. Transmission electron microscope (TEM) images of *Amoebophrya* within the cytoplasm of *Akashiwo sanguinea* during the first 10 minutes of infection prior to entering the host nucleus. A. Showing two membranes separating parasite from host (arrowheads), mitochondrion (m), small vesicles near the mitochondrion (large black arrow), nucleus (n), and nuclear envelope fibers (*) within the a nuclear envelope pocket (bar = 1 μm).

B. Showing vacuole (v), nucleus (n), mitochondrion (m), electron dense bodies (e) (bar = 1 μm). C. Showing nucleus (n), nuclear envelope fibers (*), electron dense bodies (e), electron dense body transport apparatus (white arrows) and host nucleus (hn) (bar = 1 μm). D. Showing nucleus (n), nuclear envelope fibers (*), electron dense bodies (e), and electron dense body transport apparatus (small white arrows) (bar = 1 μm).

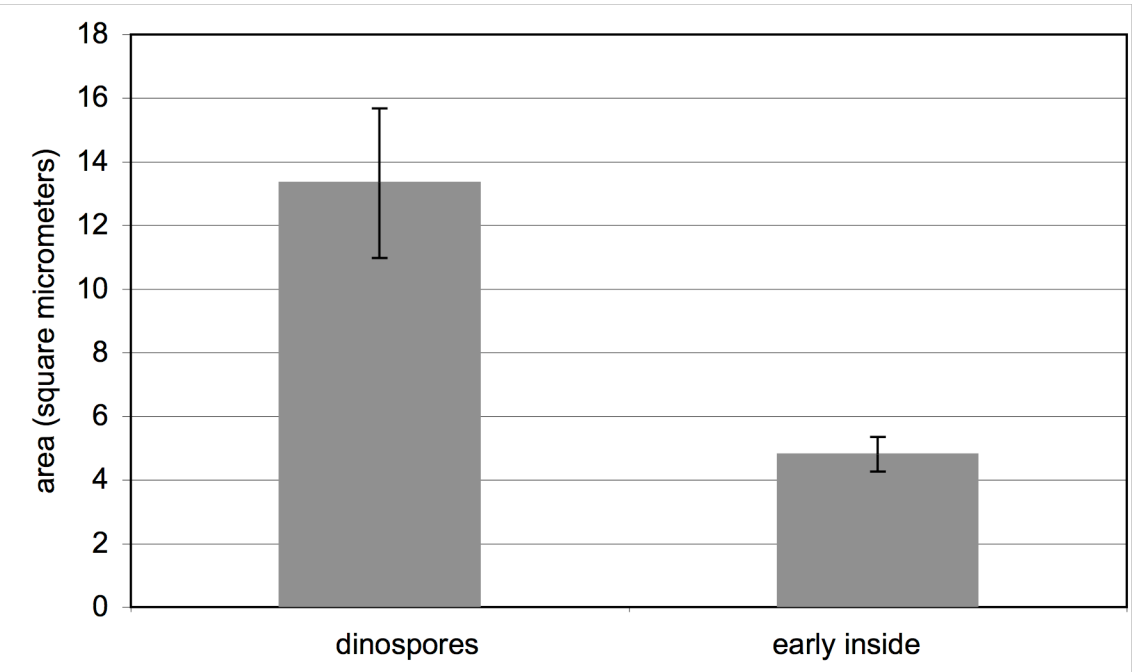


Figure 13. Shows difference in area of parasite cytoplasm in *Amoebophrya* dinospores and parasites that had been internalized for less than 10 min (Wilcoxon-Mann-Whitney: U = 1120.5, p = 1.6 x 10^{-4}).

T_{12}: At 12 hours, *Amoebophrya* is in the nucleus of *A. sanguinea* (fig 14A), where the host-parasite boundary appears to be composed of three layers (fig 14C) and in some cases is exceptionally electron dense, as in fig 14A. In this case the parasite appears to be either endocytosing host nucleoplasm or exocytosing electron dense material. A vacuole that frequently appears to be empty and associated with fibers is present in the trophont (fig 14B). Mitochondria are present at this stage but are not prominent (Fig 14B). A nucleolus is prominent (fig 14A) and the cytoplasm appears highly granular (fig 14A, B, and C) implying the presence of large numbers of ribosomes.

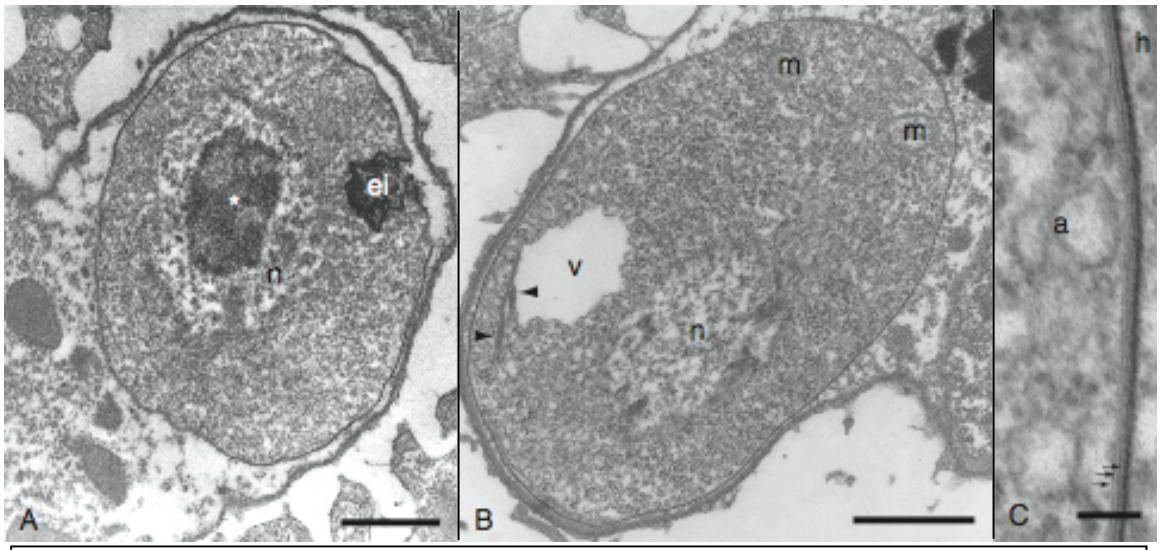


Figure 14. Transmission electron microscope (TEM) images of 12 hour *Amoebophrya* trophonts. A. Shows nucleus (n), nucleolus (*), and electron dense indentation in trophont surface (ei) (bar = 1 μm). B. Showing nucleus (n), vacuole (v), vacuole fibers (arrowheads), mitochondria (m) (bar = 1 μm). C. Showing the three layers of the trophont boundary (arrows), *Amoebophrya* cytoplasm (a), host nucleoplasm (h) (bar = 0.1 μm).

T_{24}: After 24 hours of incubation, parasite and host appeared the same as they had after 12 hours, except that the parasite was larger. One striking observation was that multiple parasites can be present in the host nucleus at this point, and they appear to interact, or at least to adhere to each other. The parasite also has a large empty vacuole and a few small mitochondria (fig 15A). They also have prominent nucleoli, but chromosomes sometimes appear to be largely decondensed (fig 15A), or else tightly condensed and closely associated with the nuclear envelope (fig 15B). Centrioles are located near the envelope of the parasite nucleus (fig 15C).

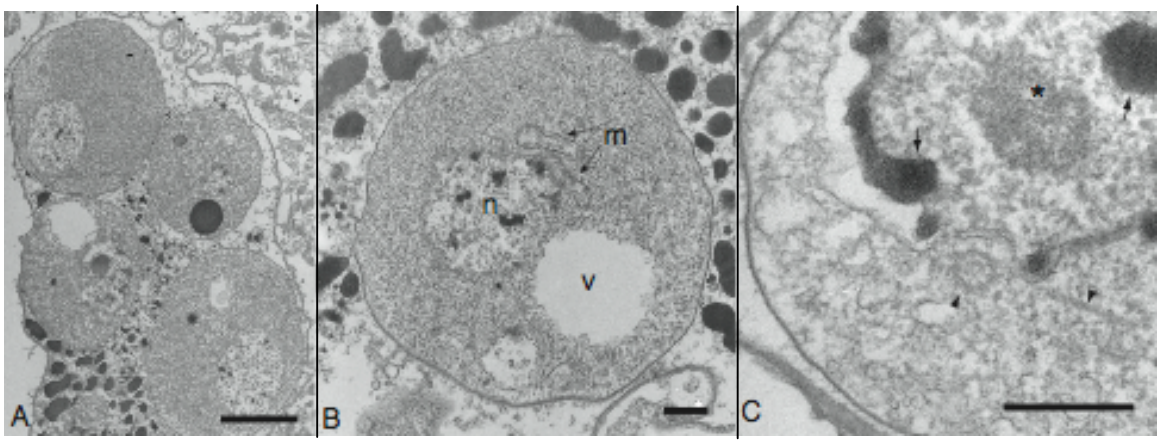


Figure 15. Transmission electron microscope (TEM) images of 24 hour *Amoebophrya* trophonts within the nucleus of *Akashiwo sanguinea*. A. Shows multiple *Amoebophrya* in a single *A. sanguinea* nucleus (bar = 5 μm). Trophont showing nucleus (n), mitochondria (m with arrows), and vacuole (v)

(bar = 1μm). C. Trophont showing nucleolus (*), condensed chromosomes (arrow), and centrioles (arrowhead) (bar = 1μm).

T_{36}: At 36 hours, the *Amoebophrya* trophont has become quite large, filling much of the host nucleus, although substantial *Akashiwo sanguinea* nucleoplasm remains. The host cytoplasm still contains a large number of plastids and lipid droplets (not shown). The cytoplasm of *Amoebophrya* remains rich with ER and small vesicles (fig 16A, B, and C). The multiple *Amoebophrya* nuclei have an irregular shape with prominent nucleoli but distinct chromosomes are absent in some trophonts. Trichocysts and striated plates have not been noted at this stage, but assembly of the mastigocoel has begun within the *Amoebophrya* cytoplasm as a narrow space associated with vesicles and basal bodies, but without flagella (fig 16C). There is a funnel lined by microtubules possibly a kind of cytopharynx (fig 16B). Near this funnel are large vesicles, which sometimes contain structures resembling host chromosomes. These vesicles are probably food vacuoles (fig 16A and B).

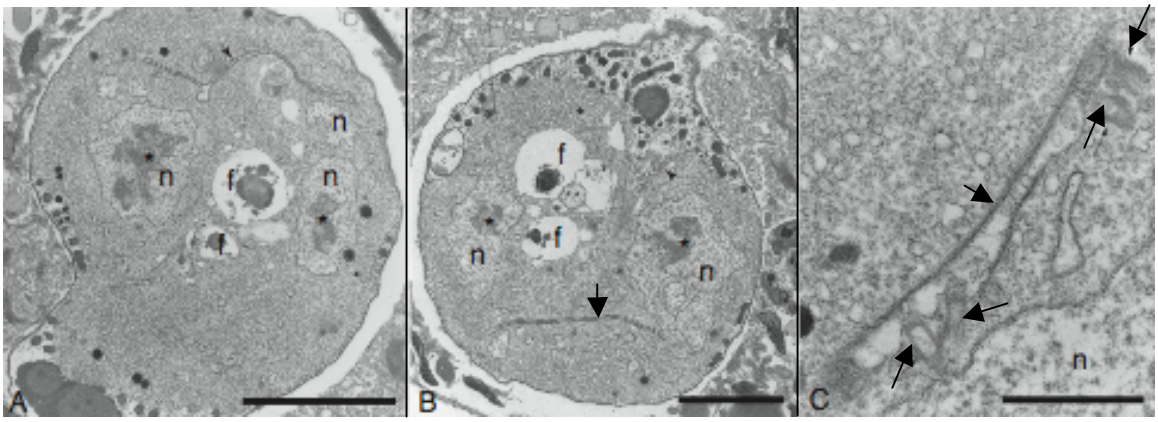


Figure 16. Transmission electron microscope (TEM) images of 36 hour *Amoebophrya* trophonts. A. nuclei (n), nucleoli (*), food vacuoles (f), and mastigocoel, (short arrows) (bar = 5μm). C. (bar = 1μm). B. Cytopharynx (c), microtubules (arrowheads), nuclei (n), nucleoli (*), food vacuoles (f), and mastigocoel (short arrows) (bar = 5μm). C. mastigocoel (short arrows), nucleus (n), basal bodies (long arrows) (bar = 1μm).

T_{48}: By 48 hours, *Amoebophrya* has largely consumed the *Akashiwo sanguinea* nucleoplasm and the host cytoplasm has been reduced to a narrow band around the parasite (fig 17A and B). *Amoebophrya* has taken on its characteristic beehive appearance with an extensively developed mastigocoel (fig 17A and B). The amphiesma-lined mastigocoel (fig 18A and B) contains numerous flagella at this stage (fig 17A, B and 18B). The flagella are arranged in rows (fig 18B) and the basal bodies show signs of being paired, have cartwheel regions (fig. 18B), and appear to be connected by a fiber at this stage (fig 18C). Multiple nuclei are present within a continuous cytoplasm and are located near

the mastigocoel (fig 17A and B). Some trophonts at this stage have largely decondensed chromatin (fig 17B) while others contain large quantities of heterochromatin (fig 17A). Within individual trophonts the level of chromatin condensation appears uniform between different nuclei. Golgi bodies, when visible, are frequently paired with nuclei (fig 19A) while mitochondria are highly elongated at this stage and one mitochondrion may pass near several nuclei (fig 19A). The central posterior region of the parasite cytoplasm is usually free of nuclei but can contain several food vacuoles (fig 17B), sometimes associated with a tube of microtutules (fig 19B). Trichocysts clustered in the *Amoebophrya* cytoplasm (fig 20A, B and C) are composed of two different portions: one has an overall circular shape and is composed of fibers and the other is more angular and striated (fig 20A, B, and C). Striated strips are frequently seen in vesicles closely associated with trichocysts or near the mastigocoel (fig 20C and D).

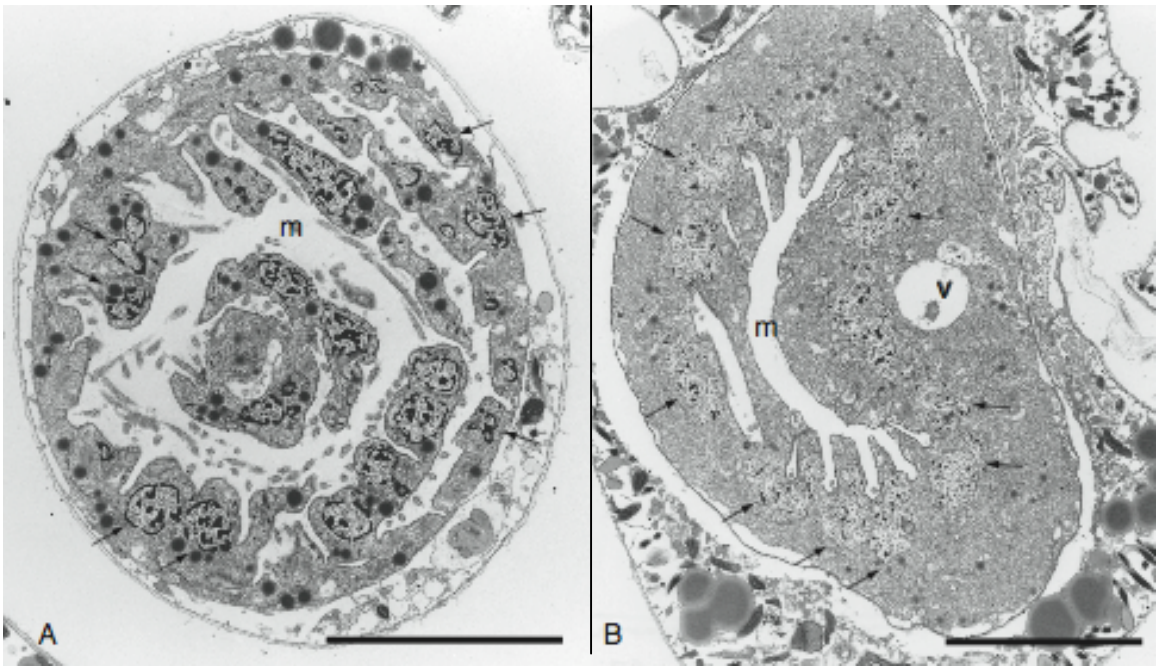


Figure 17. Transmission electron microscope (TEM) images of 48 hour *Amoebophrya* trophonts. A. Transverse section of trophont showing nuclei (arrows) and mastigocoel (m) (bar = 10µm). B. Longitudinal section showing nuclei (arrows), mastigocoel (m), and vacuole (v)(bar = 10µm).

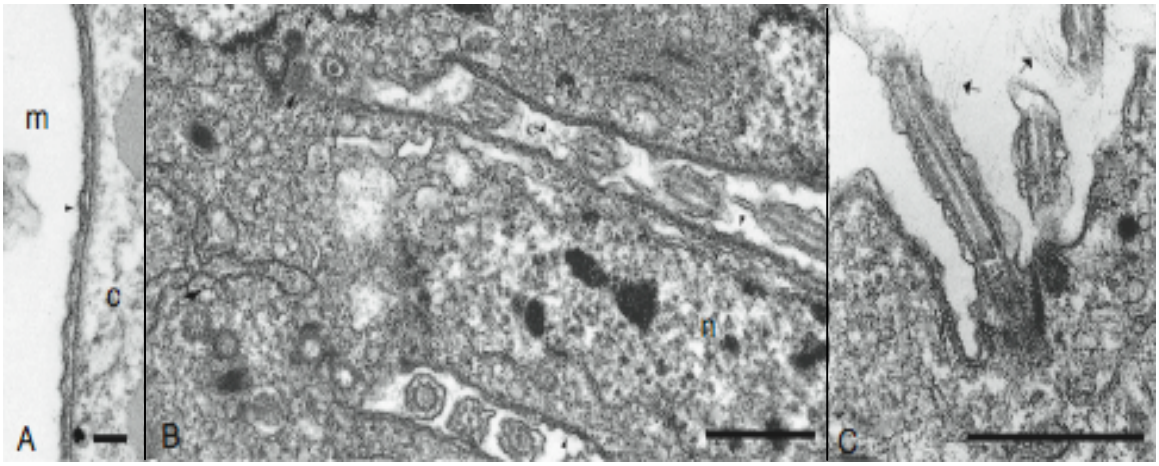


Figure 18. Transmission electron microscope (TEM) images of 48 hour *Amoebophrya* trophont. A. shows mastigocoel lumen (m), *Amoebophrya* cytoplasm (c), and amphiesma (arrowhead)(bar = 0.1 μm). B. Shows nucleus (n), amphiesma (arrowhead), and basal bodies including cartwheel regions (thick arrowhead)(bar = 1μm). C. Shows flagellum and basal body with flagellar hairs (thick arrows)(bar = 1μm).

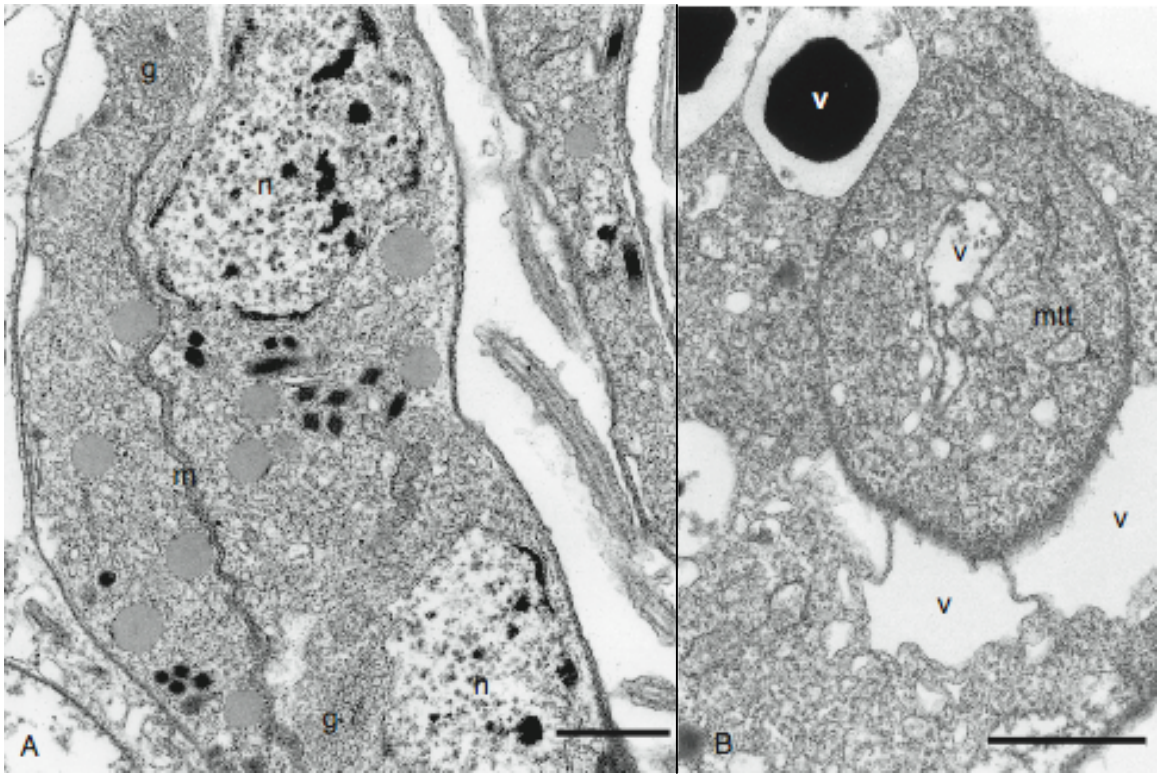


Figure 19. Transmission electron microscope (TEM) image of portions of a 48 hour *Amoebophrya* trophont. A. Shows nucleus (n), Golgi (g), and mitochondrion (m) (bar = 1µm). B. Shows microtubular tube (mtt) and food vacuoles (v) (bar = 1µm).

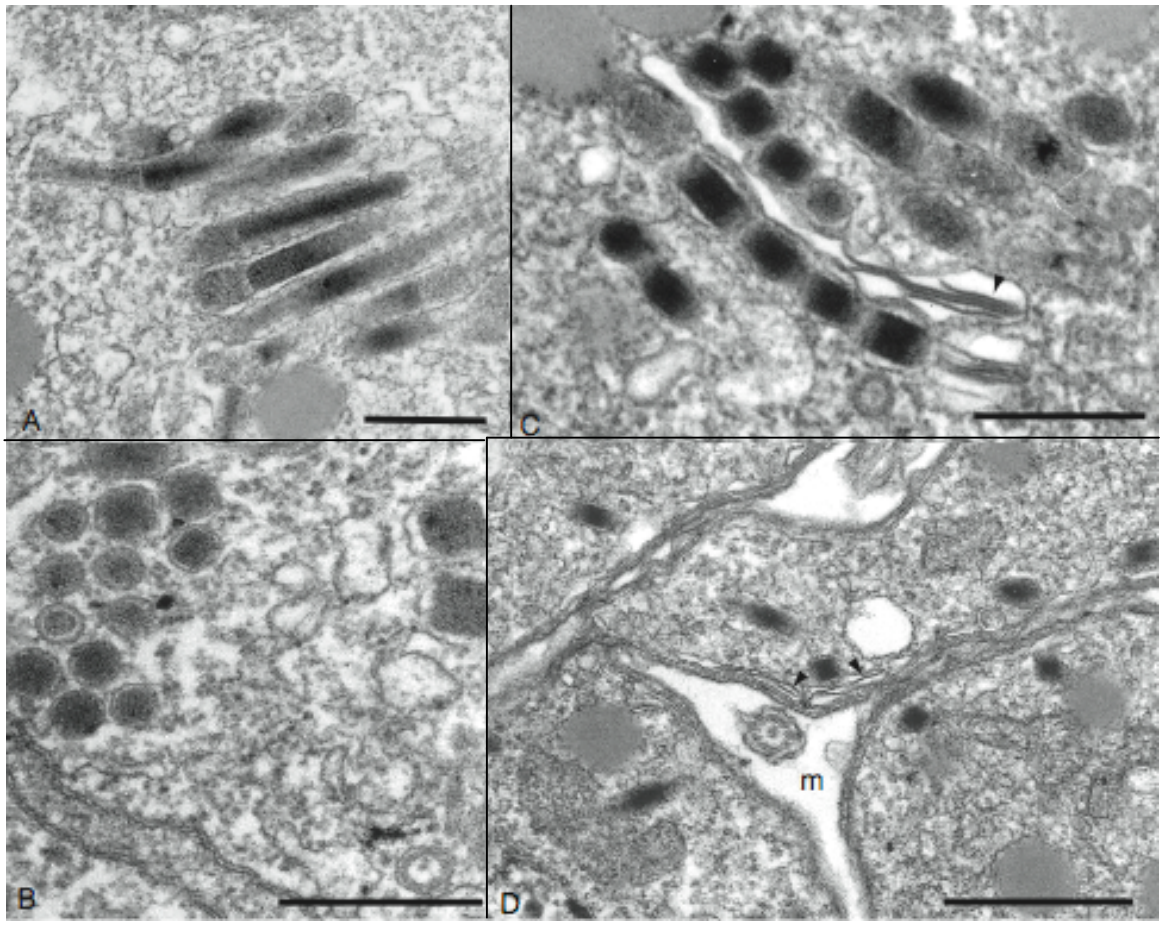


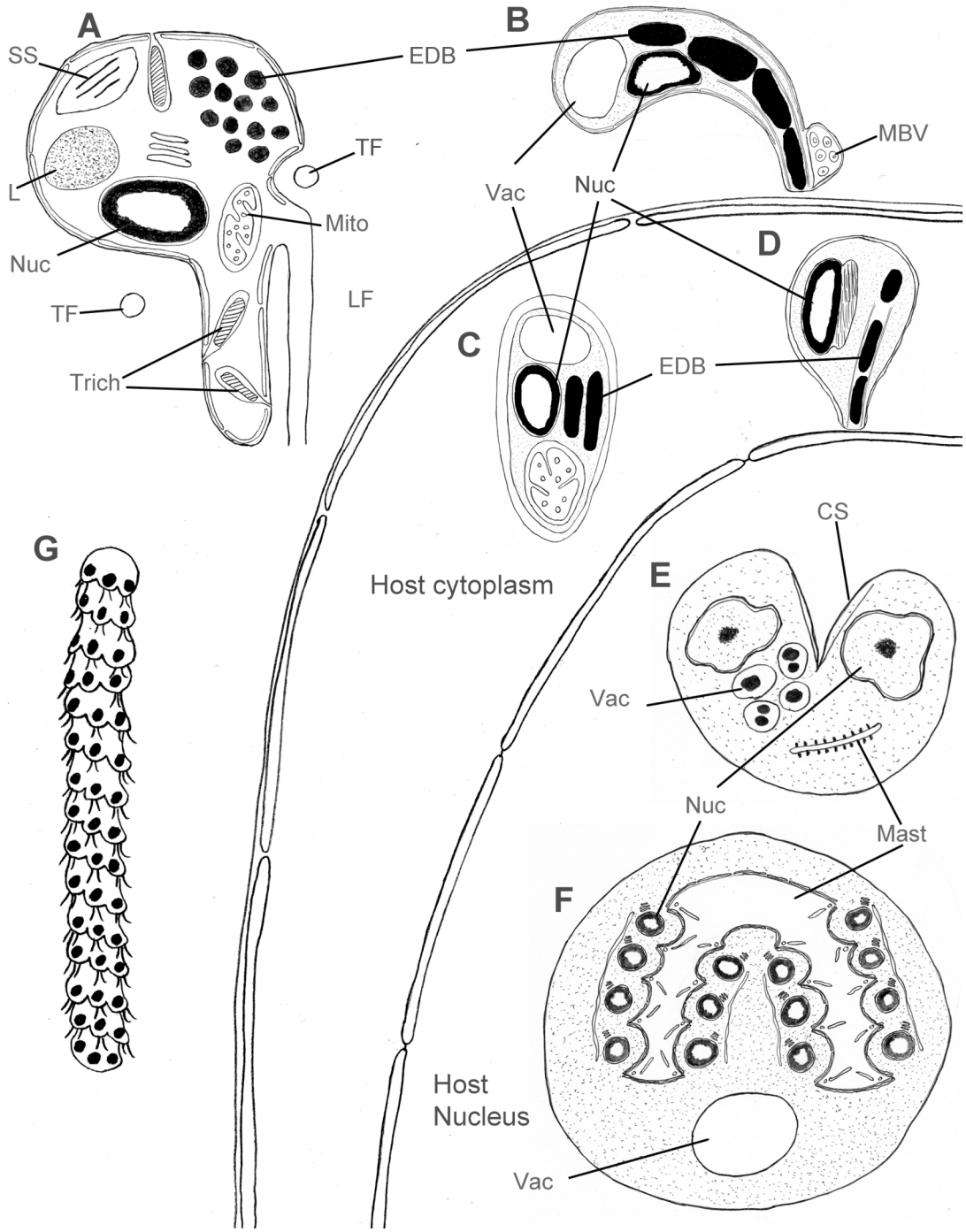
Figure 20. A) Longitudinal sections through *Amoebophrya* trichocysts showing two sections of trichocysts (0.5 µm). B) Transverse section of trichocysts showing the round distal portion and more angular proximal portion of the organelles (0.5 µm). C) Shows association of striated strips with trichocysts (0.5 µm). D) Shows mastigocoel (m), striated strips (arrowheads) (bar = 1µm).

Discussion

Finding and entering an appropriate host cell is an essential part of the life cycle of any intracellular parasite, but this stage has received relatively little attention in *Amoebophrya* research. The major stages of the life cycle of *Amoebophrya* are illustrated in figure 21. This study demonstrated that the free-swimming dinospores have trichocysts, striated strips, and electron dense bodies, which later appear in a microtubule-lined tube as the parasite enters the host cytoplasm and then the host nucleus. After attachment and penetration of the host amphiesma, the remainder of the infection process lasts somewhat more than two days. As the parasite develops within the host nucleus it forms a mastigocoel apparently through a process of vesicle fusion. Later, as the parasite takes on a beehive like appearance it grows flagella which project into the mastigocoel, lines the mastigocoel with alveoli, and synthesizes new trichocysts.

Figure 21. Shows highlights of life cycle of *Amoebophrya*. Relative sizes of stages are not to scale. A. Dinospore with Electron dense bodies (EDB), Trichocyst (Trich), Striated strips (SS), Lipid droplet (L), Nucleus (Nuc), Transverse flagellum (TF), Longitudinal flagellum (LF), and mitochondrion (Mito). B. Parasite in early stages of host entry showing Electron dense bodies (EDB), multimembrane bound vesicles (MBV) and vacuole (Vac). C. Recently entered *Amoebophrya* cell within host cytoplasm. D. Parasite entering nucleus. E. 36-hour trophont showing nascent mastigocoel (Mast)

and cytostome (CS). F. 48-hour beehive stage showing fully developed mastigocoel. G. Vermiform after leaving host. Vermiform redrawn from Cachon and Cachon 1987.



Attachment and Entry

Amoebophrya can attach to and enter *A. sanguinea* within 10 minutes of exposure. This parasite preferentially binds in the girdle or sulcus of its host or moves to these locations after binding. While it is clear from electon micrographs that *Amoebophrya* can enter its host from locations on the host surface other than the girdle or sulcus, these regions may have more recognizable surface moieties or may be easier to penetrate. While attached to the outside of its host, *Amoebophrya* forms a tube lined by microtubules encasing a series of electron dense bodies, apparently being transported along the microtubules. However, nocodazole did not inhibit entry, which implies that microtubule polymerization is not required for entry. This does not rule out a role for pre-existing microtubules. In contrast, cytochalasin D does impair entry, which implies a role for microfilament polymerization. This process has similarities to myzocytotic feeding through a peduncle by dinoflagellates. The dinoflagellate peduncle is a retractable tube lined by microtubules (Schnepf *et al*. 1985), which is sometimes associated with electon opaque vesicles (Calado *et al*. 1998; Ucko *et al*. 1997). Cytochalasins have been shown to inhibit aspects of myzocytotic feeding (Ucko *et al*. 1997; Calado *et al*. 1998). While *Amoebophrya* does not seem to have a retractable set of microtubules, its hyposome is a microtubule-lined tube. It could be using these microtubules in the entry process similar to the way *Amphidinium cryophilum* forms a feeding tube at the bottom of its hypsome (Wilcox and Wedemayer 1991). Thus there may be some evolutionary relationship between

the hypsome and the feeding tube; perhaps one started as a duplicate of the other. *Amoebophrya* does not ingest significant portions of host material until after entry.

Amoebophrya host cell entry also has important similarities to host cell entry by apicomplexans. The most obvious similarity is that cytochalasin D inhibited host cell entry by *Amoebophrya*. This indicates that *Amoebophrya* uses a microfilament-based motility system when entering a host cell, as is the case in apicomplexans (Dobrowolski and Sibley 1996; Miller *et al*. 1979; Ryning and Remington 1978; Wetzel *et al*., 2005). Although *Amoebophrya* dinospores do not have an obvious apical complex, structural evidence after binding implies that their electron dense bodies have functional similarities to apicomplexan rhoptries. The way that electron dense bodies were lined up along microtubules in attached parasites implies that the electron dense bodies are used early in the infection process but after adhesion. In apicomplexans, rhoptry contents are released after attachment during penetration (Carruthers and Sibley 1997; Dubremetz *et al*. 1998). Furthermore, it is possible that the electron dense structures, which appear to be composed of concentric layers in *Amoebophrya* dinospores, may give rise to the multi-membrane bound vesicles seen at the host-parasite boundary and the large number of vesicles within the host cytoplasm at the site of attachment. It has also been suggested that apicomplexan rhoptries are multivesicular bodies (Ngo *et al*. 2004). While inside the host cytoplasm, but prior to entering the nucleus, *Amoebophyra* is separated from the host by two membranes: their plasma membrane and a vacuole membrane. Similarly

apicomplexans introduce vesicles into their host that are rhoptry-derived and are used to form the parasitophorous vacuole membrane which protects their plasma membrane from host cell cytoplasm (Hakansson *et al*. 2001; Miller *et al*. 1979).

While structural evidence implies that the electron dense bodies are used in host cell entry and that there appear to be more electron dense bodies in free swimming dinospores compared to recently internalized parasites, my statistical analysis was not significant and this hypothesis is not supported at this time. This parasite does enter the host nucleus and again has electron dense bodies lined up along microtubules when it is in position to penetrate the nuclear envelope. *Amoebophrya* may keep leftover electron dense bodies after it has entered its host because it needs them to enter the nucleus. The parasite is significantly smaller immediately after entering its host compared to the dinospore indicating that it either uses or sheds cytoplasm that it no longer needs during the entry process.

Membrane Dynamics

As is typical of alveolates, the *Amoebophrya* dinospore has a pellicle composed of three membranes. Immediately after entering the host prior to reaching the nucleus, the parasite appears not to have alveoli, but instead to be surrounded by just two membranes: a plasma membrane and a membrane that seems to correspond to the parasitophorous vacuole seen in apicomplexans. This implies that the alveoli are lost, and possibly used, during the entry process. During its intranuclear phase, the membrane partition between *Amoebophrya* and *Akashiwo sanguinea* is composed of three layers very close together as

reported previously (Fritz and Nass 1992). It has been reported that the mastigocoel is formed by an outgrowth of the hyposome and that the mastigocoel is open in *Amoebophrya sticholonchae*, a parasite of the radiolarian *Sticholonche zanclea* (Cachon and Cachon 1987). At 48 hours, trophonts have an extensively developed mastigocoel with alveoli and numerous flagella. I interpret the narrow membrane lined spaces associated with basal bodies in 36 hour trophonts as a nascent mastigocoel. Because this structure is frequently seen associated with vesicles and I did not observe an open mastigocoel, this study indicates that the mastigocoel may form as a fully internal structure resulting from vesicle fusion in the strain of *Amoebophrya* that infects *Akashiwo sanguinea*. Other vesicles may not fuse with the mastigocoel but instead form the alveoli that line the mastigocoel in 48 hour trophonts.

Nuclear structure

Most dinoflagellates have permanently condensed chromosomes (Bhaud *et al*. 2000; Rizzo 2003). However, some organisms that diverge early from the dinoflagellate lineage do not. Among the organisms thought to be early diverging dinoflagellates, the free-swimming heterotroph, *Oxyrrhis marina*, has permanently condensed chromosomes (Dodge and Crawford 1971). *Noctiluca miliaris* (=*Noctiluca scintillans*), another free-swimming heterotroph, has a nucleus resembling that of most eukaryotes as a trophont but the chromosomes become condensed during the formation of gametes (Soyer 1970; Soyer, 1972). The syndinid parasites appear lack permanently condensed chromosomes. *Syndinium* usually appears to have some condensed chromatin, (Manier *et al*.

1971; Ris and Kubai 1974; Soyer 1974) but Soyer points out that the condensed chromatin during interphase does not correspond to chromosomes because the number of condensed units does not correspond to the number of chromosomes. Similarly, *Hematodinium* appears to have highly variable numbers of condensed chromatin units (Appleton and Vickerman 1998). *Ichthyodinium* is a group I alveolate which infects the yolk sacs of fish (Skovgaard *et al*. 2009). An organism probably in the genus *Ichthyodinium* has its chromatin condensed around the periphery of its nucleus at a late stage during its infection of cod larval yolk sacs (Pedersen *et al*. 1993). Perkinsid parasites such as *Perkinsus marinus* (=*Dermocystidium marinum)* (Perkins 1976; Perkins 1996)*, Rastriomonas subtilis* (=*Cryptophagus subtilis*) (Brugerolle, 2002; Brugerolle, 2003) and *Parvilucifera prorocentri* (Leander and Hoppenrath, 2008) have zoospores with condensed chromatin, which depending on species, may or may not appear as distinct units. The trophonts of *Syndinium* (Manier *et al*. 1971; Soyer 1974) *P. prorocentri* (Leander and Hoppenrath, 2008), and *R. subtilis* (Brugerolle, 2002; Brugerolle, 2003) sometimes have decondensed chromatin.

The degree of chromatin condensation appears to vary greatly during the life cycle of *Amoebophrya*. The *Amoebophrya* dinospore has large amounts of chromatin condensed around the periphery of its nucleus and seems to lack a nucleolus. The condensed chromatin does not appear to be divided into distinct chromosomes, but instead resembles the heterochromatin of most eukaryotic cells. Dinospores are not known to divide, and presumably have no need to synthesize DNA; transcription and translation may also be minimal. The degree

of condensation is highly variable from 12 to 48 hours and a nucleolus is prominent until 48 hours, which is consistent with the trophont growing and undergoing multiple rounds of mitosis thus requiring transcription, translation, and DNA synthesis. Trophonts usually have some level of chromatin condensation although usually less than that of the dinospores. Some trophonts were seen with no apparent condensed chromatin, however. The condensed chromatin is separated into distinct bundles at this stage. These may or may not be chromosomes given the non-correspondence of chromatin bundles with chromosomes in *Syndinium* (Soyer 1974) and *Hematodinium* (Appleton and Vickerman 1998). Despite the variability among trophonts, multiple nuclei within individual trophonts display similar levels of condensation implying a synchrony of nuclear processes. At 36 hours a nucleolus is still present, implying the trophont is still manufacturing ribosomes used in the synthesis of proteins important to dinospores (e.g., flagellar and trichocyst proteins). Some 48 hour trophonts again have large amounts of condensed chromatin lacking distinct chromosomes and some appear to lack nucleoli as is the case in dinospores. The presence of flagella and trichocysts and the lack of nucleoli implies there may be little need for additional protein synthesis at this point. These 48 hour trophonts may be preparing to sporulate. Aside from the nucleolus seen in *P. prorocentri* dinospores (Leander and Hoppenrath 2008), the nucleus of *Amoebophrya* strongly resembles that of other early diverging parasitic dinoflagellates. These parasites have highly condensed chromatin as zoospores but decondense their nuclear material during the intracellular part of their life

cycle indicating their chromatin is more like that of most eukaryotes. This suggests an evolutionary origin of permanently condensed chromosomes. The high level of chromatin condensation allows the parasite zoospores to be more compact as it looks for a new host. Perhaps dinoflagellates evolved from a parasitic ancestor, but most have lost the host associated stage of the life cycle while keeping and further modifying the highly condensed chromatin of the zoospore.

One of the more surprising things about the nucleus of *Amoebophyra* is that I did not observe a nucleus undergoing mitosis despite seeing many trophonts with multiple nuclei. This is similar to the case in *Perkinsus marinus* in which extensive efforts have been made to find mitotic *P. marinus*, yet mitosis has never been observed in this species. It has been suggested that this failure to observe mitosis in *P. marinus* is due to an extremely rapid mitosis in this species (Perkins 1969; Sunila *et al*. 2001). Perhaps the same is true of *Amoebophrya*.

Nutrition and Energy Transducing Organelles

Although it is obviously a parasite, the details of *Amoebophrya's* mode of nutrition and energy metabolism remain poorly understood. In 12 and 24 hour trophonts, electron dense invaginations of the parasite plasma membrane and empty vacuoles were present. In 36 and 48 hour trophonts, similar vacuoles contained material that appeared to be fragments of host chromosomes, implying that they are food vacuoles. The putative food vacuoles in 36 and 48 hour trophonts can sometimes be seen to be associated with a cone or tube of

microtubules that I tentatively interpret as a cytopharynx. Cachon (1964) used light microscopy to observe a fibrous channel through which material was being transported in *Amoebophrya ceratii* from *Gyrodinium* sp. Small round eosinophilic bodies were associated with this channel and Cachon referred to it as a cytopharynx. These findings are consistent with our own observations. While intracellular parasites are sometimes assumed be osmotrophic, there are other examples of intracellular parasites ingesting food by endocytosis. *Plasmodium falciparum* ring stages are actually folding over to take a “big gulp” of erythrocyte cytoplasm and later trophonts feed endocytically using a structure termed a cytostome. The *Plasmodium* cytostome appears as an electron dense indentation in the *Plasmodium* membrane similar to the electron dense indentations I observed in 12 and 24 hour *Amoebophrya* trophonts. Late stage *P. falciparum* trophonts and schizonts were observed again taking up large volumes of red blood cell cytoplasm by phagocytosis (Elliott *et al*. 2008). These large-scale phagocytic events are similar to the observation by Cachon (1964) that as vermiforms leave their host, they engulf a last large volume of host cytoplasm. Future studies could address these issues more thoroughly using modern fluorescent dyes and confocal microscopy.

The prominence of mitochondria in *Amoebophrya* cells varies greatly during the lifecycle, apparently as a function of food source. Mitochondria are easily observed in dinospores because they occupy a large portion of the cytoplasm. In 12, 24, and 36 hour trophonts they appear much smaller in relation to the cytoplasm, but by 48 hours they are highly elongated, stretching past

multiple nuclei. A similar pattern is seen in some apicomplexans. *Plasmodium falciparum* trophonts, which live in erythrocytes, frequently have mitochondria with few if any cristae but the gametocytes, which live in *Anopheles* mosquitoes, have multiple mitochondria containing more cristae (Learngaramkul *et al*. 1990). This parallels the reduction in size of mitochondria in the intracellular trophont stages of *Amoebophrya* compared to the dinospore. Within the human erythrocyte, *Plasmodium* lives in a glucose rich environment and gets most of its ATP through glycolysis (Seeber *et al*. 2008; van Dooren *et al*. 2006), but during the blood borne stages, which include merozoites, trophonts, and schizonts, starvation up-regulates mitochondrial targeted gene expression (Hall *et al*. 2005) and the sexual stages, which live in *Anopheles* mosquitoes, show greater electron transport activity (Learngaramkul *et al*. 1990). This implies that it may be using some form of oxidative phosophorylation, although its genome appears to lack several subunits of the mitochondrial ATP synthase (Gardner *et al*. 2002). From this I predict that *Amoebophrya* may use primarily glycolysis during the intracellular trophont stage, but that oxidative phosphorylation may play an important role in the dinospore. The schizonts of *Plasmodium* as well as *Toxoplasma* have highly elongated branching mitochondria with a branch associated with each nucleus prior to cytokinesis assuring that each merozoite receives a mitonchondrion (Mather and Vaidya 2008; Nishi *et al*. 2008). Similarly, the elongated mitochondria in 48 hour *Amoebophrya* trophonts may serve as a mechanism for distributing mitochondria to dinospores.

Trichocysts

Dinoflagellate trichocysts are composed of two portions: the proximal region is square in cross-section with a mesh or plate-like internal structure, and the distal portion is roundish in cross-section and composed of fibers. This is true of typical dinoflagellates such as *Gonyaulax polyedra* (=*Lingulodinium polyedrum*) (Bouck and Sweeney, 1966) as well as more basal members of the group such as *Oxyrrhis marina* (Dodge and Crawford 1971), *Noctiluca miliaris* (Sweeney 1978), and *Parvilucifera prorocentri* (Leander and Hoppenrath 2008). Dinoflagellate trichocysts usually develop in vesicles near groups of Golgi. Although layered in appearance, trichocyst primordia at times lack the sharp angular outline of the mature trichocysts (Bouck and Sweeney, 1966). The trichocysts of *Amoebophrya* appear to be similar to those typical of dinoflagellates in that they are composed of two structurally distinct regions. In 48 hour trophonts, a square cross-section was observed consistent with the proximal portion of trichocyst. Round cross-sections with subtle fibers around the periphery consistent with the distal portion of the trichocyst were also observed. Trichocyst primordia, round in cross-section but with a layered appearance, were present as well, indicating that *Amoebophrya* trichocysts and their synthesis are the same as for other dinoflagellates. *Amoebophrya* dinospores have significantly more trichocysts than the recently internalized parasites implying that if they are not used for attachment, they are shed during entry. Trichocysts then appear again in late trophonts after new ones have been synthesized.

Striated strips, which are usually found in vesicles near the periphery of dinospores, are also found in clusters of trichocysts in 48 hour trophonts suggesting these structures may be related. It could be that after *Amoebophrya* constructs trichocysts in the same manner as other dinoflagellates, it partially disassembles them into strips. This would be useful to dinospores in at least two ways: 1) The flat strips may be packed more efficiently allowing the small dinospore cell to carry more trichocysts than it could otherwise; 2) This reserve of partially disassembled trichocysts would also prevent the parasite from using all of its trichocysts at one time. Both of these factors would allow the dinospore additional opportunities to infect a host successfully.

Flagella

Dinoflagellates typically have two flagella: a longitudinal flagellum and a transverse flagellum. The longitudinal flagellum narrows at its distal end and contains a thin additional structure called a paraxial rod (Leadbeater and Dodge 1967). The typical dinoflagellate transvese flagellum has flagellar hairs and a striated strand (Leadbeater and Dodge 1967). Early diverging dinoflagellates vary from this pattern. *Noctiluca miliaris* (= *Noctiluca scintillans* Macartney) gametes have two flagella: the one considered the longitudinal flagellum has a paraxial rod, but the short transverse flagellum lacks a striated strand (Hohfeld and Melkonian 1995; Soyer, 1970). *Oxyrrhis marina* does not have a striated strand in its longitudinal flagellum but has a paraxial rod in both flagella. *O. marina* has flagellar hairs over the entire length of one flagellum and at the end of the other flagellum with narrow distal end (Dodge and Crawford 1971). The

perkinsid *Parvilucifera infectans* lacks a cartwheel region in both basal bodies and lacks some of the doublets at the end of one of its flagella (Noren *et al*. 1999). The flagella of *Amoebophrya* most closely resemble those of typical dinoflagellates. The longitudinal flagellum narrows at its distal end and contains an accessory structure that I interpret as being homologous to a paraxial rod. Its transverse flagellum has flagellar hairs and encircles the dinospore beginning in an indentation, which could be interpreted as a girdle. It contains an accessory structure, which could be homologous to a striated strand, although it is much less obvious than the striated strand of most dinoflagellates and in that way seems more similar to the second paraxial rod of *O. marina*. *Amoebophrya* does have a cartwheel region in its basal bodies and no flagella lacking doublets were observed, distinguishing it from *P. infectans*. A swelling of the nuclear envelope of *Amoebophrya* sometimes contains fine fibrillar material. Studies of other organisms including the dinoflagellates *Hematodinium* (Appleton and Vickerman, 1998), *Oxyrrhis marina* (Dodge and Crawford 1971), and *Woloszynskia micra* (Leadbeater 1971), as well as other algae (Leedale *et al*. 1970) have indicated that this material is composed of flagellar hairs.

Evolution of parasitism in Alveolata

Phylogenetic studies have indicated that dinoflagellates and apicomplexans are closely related (Fast *et al*. 2002; Gajadhar *et al*. 1991; Wolters 1991). *Amobeophyra* host cell penetration resembles both host cell penetration in apicomplexans and myzocytosis in dinoflagellates. The longitudinal microtubules in its hyposome resemble both apicomplexan cortical

longitudinal microtubules and dinoflagellate peduncle microtubules. The electron dense vesicles of *Amoebophrya* show similarity to apicomplexan rhoptries and the electron dense vesicles associated with dinoflagellate peduncles. Myzocytosis as well as host cell entry by apicomplexans and *Amoebophrya* are inhibited by cytochalasins.

There are examples of parasitism spread throughout the dinoflagellate lineage. *Perkinsus* (Leander and Keeling 2004), *Syndinium*, and *Amoebophrya* (Skovgaard *et al* 2005) diverge early but several parasitic organisms including *Amyloodinium ocellatum* (Litaker *et al* 1999), *Blastodinium* spp. (Skovgaard *et al*. 2007), *Paulsenella vonstoschii* (Kühn and Medlin 2005), and *Piscinoodinium* sp. (Levy *et al*. 2007) are nested within the mostly non-parasitic dinokaryota. *Colpodella* is a myzocytotic feeder at the base of the apicomplexan lineage (Brugerolle 2003; Leander *et al* 2003; Myl'nikov 2000; Simpson and Patterson 1996). One obvious explanation for the similarity of entry mechanisms between *Amoebophrya* and apicomplexans would be common ancestry, but this would imply that the common ancestor was a parasite. If this hypothesis were correct, it would be necessary to explain why *Amoebophrya* would retain the life style of an intracellular parasite and a similar mechanism of entry but lose its apical complex. This might be explained if the common ancestor was a parasite with a complex life cycle which included a capable free-swimming stage and that the dinoflagellates and apicomplexans specialized in different parts of the life cycle thus taking different characteristics away from the common ancestor. This model would imply that the common ancestor of the apicomplexan-dinoflagellate clade

was more *Amoebophrya*- or *Perkinsus*-like than *Colpodella*-like because *Colpodella* lacks a parasitic stage, although it has an apical complex. This implies that the apicomplexans may be rooted among the early diverging dinoflagellates. *Chromera* also has a life cycle that includes free-living and non-obligate symbiotic stages which is perhaps more compatible with the above model, but if the last common ancestor of the apicomplexan-dinoflagellate clade was more *Chromera*-like, the dinoflagellates would be rooted among the early diverging apicomplexans. In this case, *Amoebophrya* might have split from the apicomplexans prior to the apicomplexans acquiring an apical complex. A second possibility to explain the similarity in entry mechanisms would be that this is an example of convergent evolution and that *Amoebophrya* evolved an apicomplexan-like mechanism of entry because it is an effective mechanism of entering a host cell. In this case, *Amoebophrya* and *Perkinsus* might share a most recent common ancestor that was not parasitic.

Conclusions

Amoebophrya shares many characteristics with other dinoflagellates. It has two flagella, both of which have accessory structures and one of which has flagellar hairs. It also has an arsenal of dinoflagellate style trichocysts. *Amoebophrya* also shares several characteristics with apicomplexans. It is an intracellular parasite, which uses a microfilament based motility to enter host cells and it has electron dense organelles that, while they are not docked at an apical complex, they do appear to have structural and functional similarities to rhoptries. Given the wide variety of similarities and differences among early

diverging dinoflagellates and apicomplexans, a new hypothesis regarding the origins of these groups was proposed. The common ancestor of the dinoflagellates and apicomplexans may have been a parasite with a complex life-cycle that included a swimming stage similar to *Perkinsus* or *Amoebophrya*.

Chapter 3: *Emiliania huxleyi* Phylogenomics

Abstract

Emiliania huxleyi is a haptophyte alga. Haptophytes are important members of marine phytoplankton and play vital roles in global biogeochemistry. Haptophytes possess red algae derived plastids, as do cryptophytes and the plastidic members of the heterokonts and alveolates. This similarity resulted in these groups being put into the chromalveolate supergroup. This study used the predicted proteins from the complete *E. huxleyi* genome to do a phylogenomic analysis. The analysis revealed that *E. huxleyi* has genes associated with several lineages indicating that haptophytes are deeply chimeric organisms. The strongest signals were from heterokonts, green algae, and red algae. The most likely interpretation is the haptophytes are closely related to heterokonts and have received genes from the red algal endosymbiont. The large green algal input is more difficult to interpret. It could be that all or some portion of the chromalveolates had a green algal plastid before they had a red algal derived plastid or that chromalveolates are actually descended from green algae. However, there is a lot more genomic data available from the Viridiplantae (=green lineage) than from the rhodophytes (=red lineage) so it is also possible that the green algal signal is part of a primary plastid lineage signal and as more red algal genomes become available the apparent green algal signal will decrease to insignificance.

Introduction

Haptophyte algae are among the three most abundant groups of eukaryotic phytoplankton in marine ecosystems, the other major lineages being diatoms and dinoflagellates (Falkowski *et al*. 2004). Precise estimates of carbon fixation are difficult to obtain, but they are clearly among the most important primary producers in marine environments. In addition to primary productivity, they can play a role in the long-term sequestration of carbon because they mineralize $CaCO_3$ during the formation of coccoliths. The biomineralization of $CaCO_3$ produces CO_2, but the activity of haptophytes has been shown to result in a net drawdown of carbon from the atmosphere as a result of their photosynthetic activity (van der Wal *et al*. 1995). Biomineralization by these microbes is responsible for much of the chalk in cretaceous sediments and large geologic structures like the White Cliffs of Dover (Mitchell *et al*. 1997). Because they are easily cultured, they have been used as a model organism to study the cell biology (Marsh 2003) and ecology (Iglesias-Rodriquez *et al*. 2008; Riesbesell *et al*. 2000) of $CaCO_3$ biomineralization. This is of particular concern at present because of the potential impact of rising CO_2 levels and climate change on $CaCO_3$ mineralizing organisms including reef-building corals.

Haptophytes also are involved in the sulfur cycle because they release dimethylsulphoniopropionate (DMSP), a major component of atmospheric sulfur. DMSP is, in turn, converted to dimethyl sulfide (DMS), which promotes cloud condensation and thus leads to an increase in albedo, which has the effect of

reducing the warming effects of the sun (Charlson *et al*. 1987). The DMS is oxidized to sulfuric acid, which in turn can lead to acid rain. It has been hypothesized that the large numbers of haptophytes during the Cretaceous resulted in acid rain that lead to high levels of extinction prior to the K–T boundary when most remaining dinosaurs went extinct (Robinson 1995).

The history of the systematics of haptophytes and heterokonts is perhaps almost as complex as the evolutionary history of these organisms. Most organisms now considered haptophytes were put in the class Chrysophyceae by Pascher in 1914 along with organisms still considered chrysophyte heterokonts. Christensen (1962) first named this group the class Haptophyceae. The class Chrysophyceae included organisms with a golden-brown color and one or two flagella, which could be of equal or unequal length. The chrysophytes were later split into four groups by Bourrelly (1965). One of these groups was the Isochrysophycidae, a group that overlaps greatly with the modern haptophytes. Hibberd (1976) formally removed the haptophytes from the Chrysophyceae, pointing out that true chrysophytes share more ultrastructural features with the Bacillariophyta (diatoms), xanthophytes and the phaeophytes (include giant kelp) than with the haptophytes. The latter he placed in the class Prymnesiophyceae, which he equated with the Haptophyceae. Haptophytes share a variety of features with heterokonts including the presence of chlorophylls a and c, plastids surrounded by four membranes with thylakoids in stacks of three, and storage of chrysolaminarin (Graham and Wilcox 2000). The main feature that unifies the heterokonts and that distinguishes them from haptophytes is the presence in

heterokonts of tripartite mastigonemes on one flagellum (Anderson 2004). Bhattacharya *et al*. (1992) found that haptophytes fell outside the heterokonts based on SSU rDNA. Later, Bhattacharya *et al*. (1993) using a broader taxon sampling and actin sequences, placed *Emiliania huxleyi* between animals and a group of amoebae including *Dictyostelium discoideum*. More recently several studies using sequences from multiple genes have placed haptophytes as sister to cryptomonads (Patron *et al*. 2007) and put the rhizaria and the alveolates closer to the heterokonts (Burki *et al*. 2007; Hackett *et al*. 2007). Furthermore, presence of a significant green algal signal in the genomes of heterokonts and haptophytes has recently been proposed (Moustafa *et al*. 2009).

The diatoms as well as other heterokonts and the haptophytes were placed in the group Chromophyta by Cavalier-Smith (1981) largely because they have plastids that contain chlorophylls a and c and a stroma surrounded by four membranes rather than the two found in red and green algae. The heterotrophic heterokonts were included because they have flagellar mastigonemes similar to those of flagellate plastidic heterokonts. The cryptophytes were excluded because they had phycobilins and a nucleomorph associated with their plastids. Later, Cavalier-Smith (1999) grouped the alveolates and the cryptophytes with the chromists and called them chromalveolates. This put all of the chlorophyll c containing organisms in one group despite most dinoflagellates having only three membranes around their stroma and ciliates not having plastids. It is widely believed that chlorophyll c containing plastids are descended from engulfed rhodophytes, which became mutualistic symbionts and later organelles. Thus,

these organisms should contain genes of rhodophyte origin. This is the case with the diatoms. *Phaeodactylum tricornutum* shares at least 76 genes with green/red algae and 490 with the green/red lineage and the chromalveolates to the exclusion of other taxa. *Thalassiosira pseudonana* shares at least 180 genes with the green/red algal lineage and 539 genes with the green/red lineage and the chromalveolates to the exclusion of other taxa (Bowler *et al*. 2008). For these reasons, it is expected that *Emiliania* will have genes from the red/green lineage.

Conventional phylogenomic methods like supermatrix and supertree methods average the signal from a lot of genes. Edwards (2008) pointed out that these strategies are limited because they miss such phenomena as deep coalescence, horizontal gene transfer, and paralogy. These techniques are therefore inappropriate for the study of chimeric organisms like *E. huxleyi*. One promising approach to this problem is the concaterpillar package (Leigh *et al*. 2008), which uses a likelihood-model based approach to test for phylogenetic combinability of data partitions. Unfortunately, this program is computationally intensive and consequently difficult to apply to very large data sets. For these reasons I chose to take a different approach to the analysis of discordant phylogenetic signals, and designed a phylogenomic pipeline that sorts individual gene trees based on the clades in which the study taxon (in this case *E. huxleyi*) occurs.

Methods

Pipeline overview

A pipeline was designed to find orthologous sequences, produce individual protein trees for each COG, count clade frequencies, sort trees according to what clade they support, and find concatenated trees. Figure 1 is a diagram of the pipeline, which is explained in more detail below.

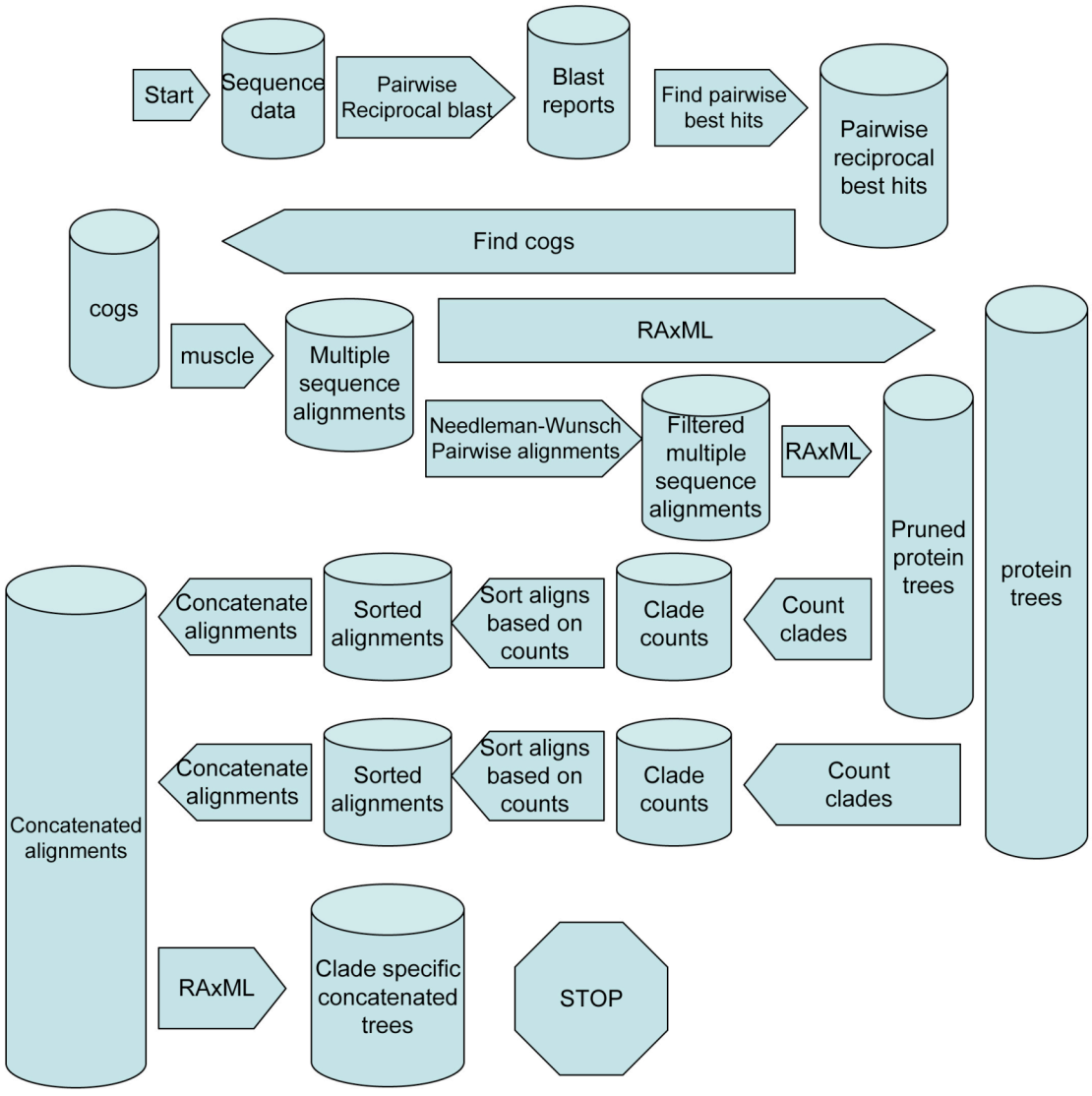


Figure 1: Diagram of phylogenomic pipeline

Taxon sampling

A list of complete genomes was chosen so that each major lineage was represented by three or four genomes. Additional genomes and EST libraries for the heterokonts, rhizaria, viridiplantae, and rhodophyta were added later as they were deemed particularly important based on the literature and preliminary analyses. A complete list of organisms as well as data sources can be found in table 1.

Finding orthologous sequences

Predicted proteins for complete genomes were downloaded from NCBI, JGI, and EuPathDB. EST projects were assembled into contigs using CAP3 (Huang and Madan 1999) then translated in all six reading frames using transeq from the EMBOSS suite (Rice *et al*. 2000). A PERL script was written that serially BLASTed (Altschul *et al*. 1997) predicted protiens from each organism against those from each other organism in the database, found pairwise reciprocal best hits (PRBH) using the bioperl Bio::SearchIO module (Stajich *et al*. 2002), and found clusters of orthologous groups (COGs) (Tatusov *et al*. 1997) from the PRBH. COGs were found by creating three duplicate arrays of PRBH and finding combinations of three pairs such that each pair shared one taxon with each of the other two. Because each pair could have this relationship with multiple combinations of other pairs, large groups of sequences were sometimes found. COGS were aligned using MUSCLE (Edgar 2004). Alignments were viewed in Maclade or Jalview or as text files.

Organism	Data	Source	Citation
Alexandrium_tamarense	est	NCBI	Hackett *et al*. 2005
Arabidopsis_thaliana	genomic	NCBI	The Arabidopsis Genome Initiative
Aspergillus_fumigatus	genomic	NCBI	Nierman *et al*. 2005
Aureococcus_anophagefferens	est	NCBI	Richardson *et al*. unpublished
Bigelowiella_natans	est	NCBI	Keeling unpublished
Caenorhabditis_elegans	genomic	NCBI	The *C. elegans* Seq Cons 1998
Chlamydomonas_reinhardtii	genomic	JGI	Merchant *et al*. 2007
Cryptosporidium_parvum	genomic	EuPathDB	Abrahamsen *et al*. 2004
Cyanidioschyzon_merolae	genomic	NIG, Japan	Matsuzaki *et al*. 2004
Dictyostelium_discoidium	genomic	NCBI	Eichinger *et al*. 2005
Drosophila_melanogaster	genomic	NCBI	Adams *et al*. 2000
Ectocarpus_siliculosus	est	NCBI	Dittami *et al*. 2009
Emiliania_huxleyi	genomic	JGI	Read et al. unpublished
Entamoeba_histolytica	genomic	NCBI	Loftus *et al*. 2005
Escherichia_coli_K12_MG1655	genomic	NCBI	Blattner *et al*. 1997
Euglena_gracilis	est	NCBI	Grey unpublished
Fucus_serratus	est	NCBI	Peason *et al*. 2010
Giardia_lamblia	genomic	EuPathDB	Morrison *et al*. 2007
Guillardia_theta	est	NCBI	Lucas *et al*. unpublished
Guillardia_theta_nucleomorph	genomic	NCBI	Douglas *et al*. 2001
Heterocapsa_triquetra	est	NCBI	Patron *et al*. 2005
Homo_sapiens	genomic	NCBI	Int Hum Genome Seq Cons 2001
Karenia_brevis	est	NCBI	Lidie *et al*. 2005
Leishmania_infantum	genomic	NCBI	Peacock *et al*. 2007
Oryza_sativa (japonica)	genomic	NCBI	Goff *et al*. 2002
Ostreococcus_tauri	genomic	JGI	Derelle *et al*. 2006
Paracercomonas_marina	est	NCBI	Rodriquez-Ezpeleta et al. 2007
Paramecium_tetraurelia	genomic	ParameciumDB	Aury *et al*. 2006
Phaeodactylum_trichornutum	genomic	JGI	Bowler *et al*. 2008
Physcomitrella_patens	genomic	JGI	Rensing et al. 2008
Phytophthora_ramorum	genomic	JGI	Tyler et al. 2006
Phytophthora_sojae	genomic	JGI	Tyler et al. 2006
Plasmodium_falciparum	genomic	EuPathDB	Gardner et al. 2002
Porphyra_haitanensis	est	NCBI	Shen *et al*. unpublished
Porphyra_yezoensis	est	NCBI	Shen *et al*. unpublished
Reticulomyxa_filosa	est	NCBI	Burki et al. 2006
Saccharomyces_cerevisiae	genomic	NCBI	Goffeau et al. 1996
Schizosaccharomyces_pombe	genomic	NCBI	Wood et al. 2002
Strongylocentrotus_purpuratus	genomic	NCBI	Sea Urchin Gen Seq Cons 2006
Synechocystis_6803	genomic	NCBI	Kaneko et al. 1996
Thalassiosira_pseudonana	genomic	JGI	Armbrust et al. 2004
Toxoplasma_gondii	genomic	EuPathDB	Gajria et al. 2007
Trichomonas_vaginalis	genomic	EuPathDB	Carlton et al. 2007
Trypanosoma_brucei	genomic	NCBI	El-Sayed et al. 2005

Table 1: List of genomes used and sources.

Because of concern about results being due to poorly aligned sequences, the Needleman-Wunsch algorithm was used to remove sequences with scores below 800 (Needleman and Wunsch 1970). This global alignment procedure was used because it was assumed that COGs that had been aligned should represent proteins and should be alignable from end to end. A PERL script was written that used needle from the EMBOSS package (Rice *et al*. 2000) to align each sequence in a COG to the *Emilinia huxleyi* sequence in that COG (gap opennig penalty 10.0; gap extension penalty 0.5). Sequences with scores below 800 were removed and analyses were repeated with the cleaner data. The filtration cutoff of 800 was the score that kept the maximum number of sequences while eliminating poorly aligned sequences as determined by visual inspection. Needleman-Wunsch scores were also used to count the two most alignable sequences to the *E. huxleyi* sequence within each COG.

Phylogenetic analysis of individual alignments

Maximum likelihood trees were found for each amino acid alignment using RAxML with a WAG amino acid replacement matrix (Whelan and Goldman 2001) and a gamma distribution was used to model substitution rate variation (Stamatakis 2006). New trees were calculated in the same manner for alignments that had had sequences with Needleman-Wunsch scores below 800 removed. RAxML did 100 bootstrap trees for each individual alignment using its rapid bootstrapping method. Resulting trees were viewed with fig tree. A PERL script was written that extracts and counts all bipartitions from each tree

containing the exact same taxa. Bipartitions with exactly the same taxa content were initially counted by a PERL script that split each tree at each parenthesis going one direction creating a list of all possible splits then for each fragment all possible splits were made based on parentheses going the other way. The fragments with balanced left and right parentheses were then colleted as bipartitions and those with exactly the same taxa were counted. The bipartitions found by the script for individual trees were compared manually to corresponding trees to confirm the script was correctly identifying bipartitions. Tree specific files containing lists of bipartitions were used to search for and count bipartitions that represent subsets of the organisms in a lineage using another PERL script I wrote.

Concatenated alignments

A PERL script was written that concatenated protein alignments by progressively adding sequences from individual protein alignments for each organism. If an organism lacked a particular protein sequence, gap characters were inserted to maintain the alignment. Paralogous sequences were removed to maintain the alignment. Organisms that did not contribute and sequence to a concatenated alignment were removed.

Best trees for most concatenated alignments were found with RAxML rapid bootstrapping (f –a option)(Stamatakis 2006). All analyses were run with the JTT (Jones *et al*. 1992) matrix and most were also run with the WAG matrix (Whelan and Goldman 2001). A gamma distribution was used to model the substitution rates. In the two largest concatenated analyses, those linking

Emiliania huxleyi to the heterokonts and then separately to the green lineage, the analyses did not finish and no best trees were found. In these two cases, the f –b option was used to find the best trees. Ten replicate trees were calculated in each case and two trees from each analysis that seemed representative of the other trees were chosen. Because the bootstrap files in these two cases did not contain 100 bootstrap trees, two corresponding bootstrap files from separate incomplete runs for each of these two analyses were combined. The total number of bootstrap replicates used for each analysis is shown in table 2. A PERL script was written to extract numberical bootstrap values from the bootstrap files from the incomplete f –a runs. Extended majority rule consensus trees were found using consense from the Phylip suite (Felsenstein 1989). The bootstrap values from the consensus trees and the counts from the PERL script were compared and were found to be very similar although the best trees from the RAxML f –b runs did not have exactly the same topologies as the consensus trees so there was not an exact correspondence. The best tree from each f –b run with the highest average bootstrap values was used as the best tree. Bootstrap values were manually applied to each of the two best trees produced with the f –b option in Adobe Photoshop Elements (Adobe Systems Incorporated).

Required bipartition	proteins	amino acids	taxa	bootstraps
E.huxleyi heterokont	305	294567	41	166
E.huxleyi heterokont reduced		268427	41	
E.huxleyi heterokont 800	115	127621	37	105
E.huxleyi heterokont 800 reduced		108320	37	
E.huxleyi greens	180	212322	38	192
E.huxleyi greens reduced		201703	38	
E.huxleyi greens 800	87	118374	37	105
E.huxleyi greens 800 reduced		104150	37	
E.huxleyi reds	46	47122	38	105
E.huxleyi reds reduced		44684	38	
E.huxleyi reds 800	29	36321	36	105
E.huxleyi reds 800 reduced		31025	36	
Alveolates heterokonts	55	64459	36	105
Alveolates heterokonts reduced		41178	36	
Alveolates heterokonts 800	44	64459	37	105
Alveolates heterokonts 800 reduced		47812	37	
E.huxleyi cryptomonad	4	2691	31	105
E.huxleyi cryptomonad reduced		2533	31	
E.huxleyi cryptomonad 800	1	1245	29	105
E.huxleyi cryptomonad 800 reduced		1244	29	
E.huxleyi heterokonts greens	67	66121	37	105
E.huxleyi heterokonts greens reduced		63115	37	
E.huxleyi heterokonts greens 800	19	30783	36	105
E.huxleyi heterokonts greens 800 reduced		30183	36	
E.huxleyi heterokonts reds	52	49452	36	105
E.huxleyi heterokonts reds reduced		46478	36	
E.huxleyi heterokonts reds 800	14	14531	31	105
E.huxleyi heterokonts reds 800 reduced		13263	31	

Table 2: Number of proteins, amino acid residues, taxa, and bootstrap replicates used in concatenated analyses.

To find out if the results for each concatenated tree were repeatable, samples were taken without replacement from concatenated alignments using the sample function in R (www.Rproject.org). The initial sample was 80,000 residues long. A perl script was used to take the columns 4000 at a time and put them into 20 phylip formated alignments. Best trees for subsampled alignments were found using RaxML.

Gene Ontologies

GeneMerge (Castillo-Davis and Hartl 2003) was used to identify functional trends in genes grouping *E. huxleyi* to various bipartitions. *Arabidopsis* specific gene association and description files were constructed using specially written PERL scripts based on files downloaded from the TAIR website (ftp://ftp.arabidopsis.org/home/tair). Protein accession numbers were converted to gene locus tags using the TAIR9_NCBI_REFSEQ_mapping_PROT file. Gene association and description files were made using information in the ATH_GO_GOSLIM.txt file. The population file was based on all of the predicted protiens in the *A. thaliana* genome. Study files were lists of all the *A. thaliana* genes found in COGs that produced trees containing particular clades. Resulting analyses were based on fewer genes than the concatenated alignments because *A. thaliana* was not present in every individual protein alignment and not all *A. thaliana* sequences were associated with GO terms.

Synteny test

To test for synteny, *E. huxleyi* protein sequences from COGs clustered by the bipartition within which *E. huxleyi* fell out were blasted against the scaffolds of the *E. huxleyi* genome using tblastn. A PERL script was written to quantify best hits per scaffold. A Fisher exact test was run using R (www.Rproject.org) to test the null hypotheses that hits per scaffold would be similar to hits predicted by scaffold length.

Results

The pairwise reciprocal best-hit blast search resulted in 1656 non-redundant COGs which contained *Emiliania huxleyi*. Phylogenetic trees were found for only 1129 of these COGs because MUSCLE was unable to find alignments or RaxML was unable to find trees for some COGs. The COGs and thus the trees did not all contain the same taxa and taxa represented by complete genomes were in more trees than those represented by EST libraries (fig 2). The taxonomic content of the COGs for which trees were made is shown in the Appendix. When bipartitions containing *E. huxleyi* and exactly the same set of taxa were quantified, this haptophyte fell out most frequently with diatoms and second most frequently with the rhodophyte *Cyanidioschyzon merolae*, but it also frequently fell out with non-diatom heterokonts and members of the Viridiplantae (fig 3). *Emiliania* grouped most often with heterokonts in general and more frequently with green lineage organisms in general than with red lineage organisms (fig 4). In this study, Alveolates grouped with heterokonts to the exclusion of *E. huxleyi* 55 times and *E. huxleyi* only grouped with cryptophytes (=cryptomonads) in four protein trees compared to 305 protein trees in which *E. huxleyi* fell out with the heterokonts and 180 with green lineage organisms. *Emiliania* grouped with the heterokonts and reds together 52 times, surprisingly close to the 67 times *E. huxleyi* grouped with heterokonts and greens together given the large discrepancy between the number of times *E. huxleyi* grouped with any group of greens compared to *E. huxleyi* grouping with any red.

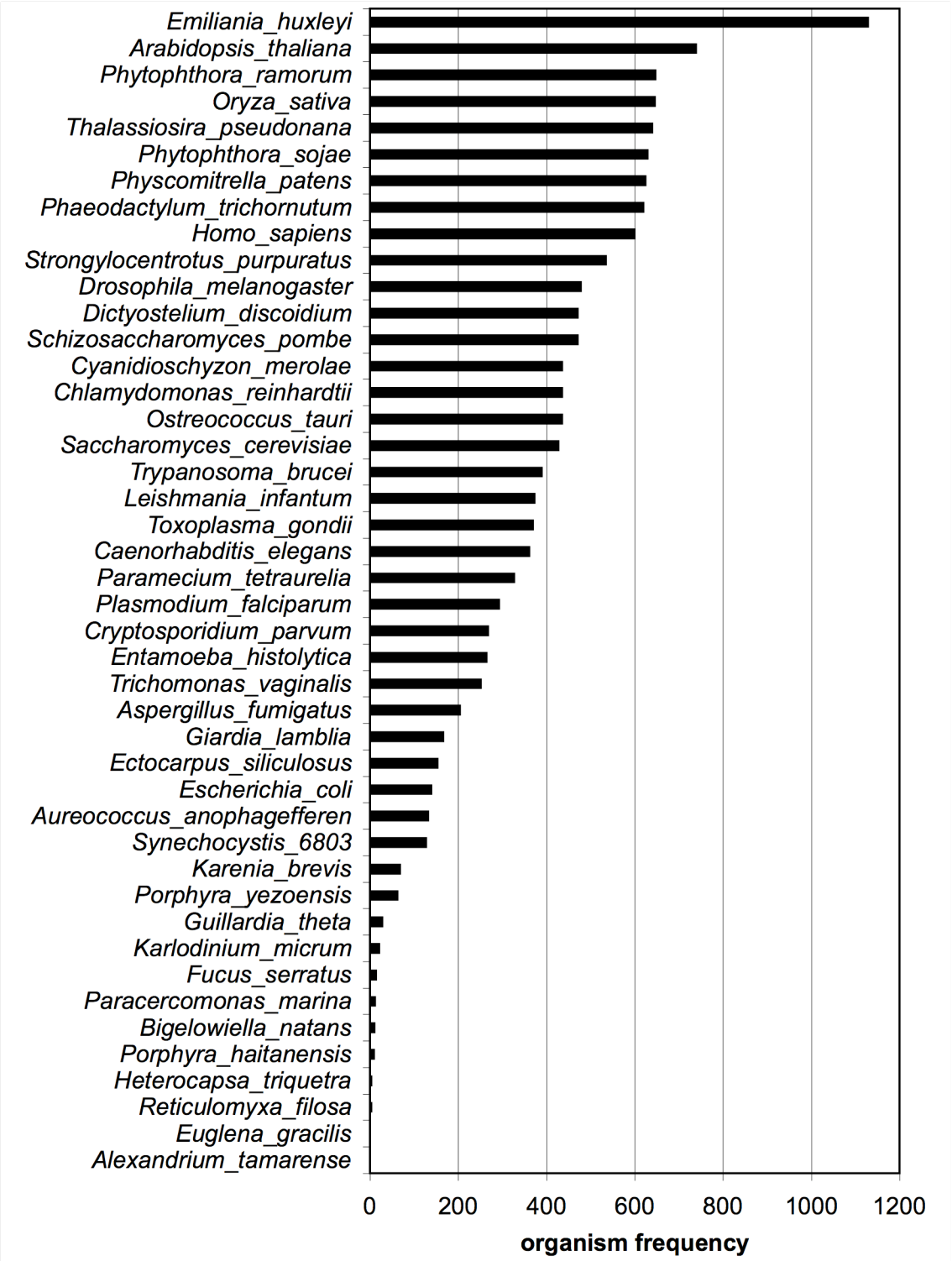


Figure 2: Frequencies of organisms in 1129 *Emiliania huxleyi* containing COGs for which individual protein trees were found.

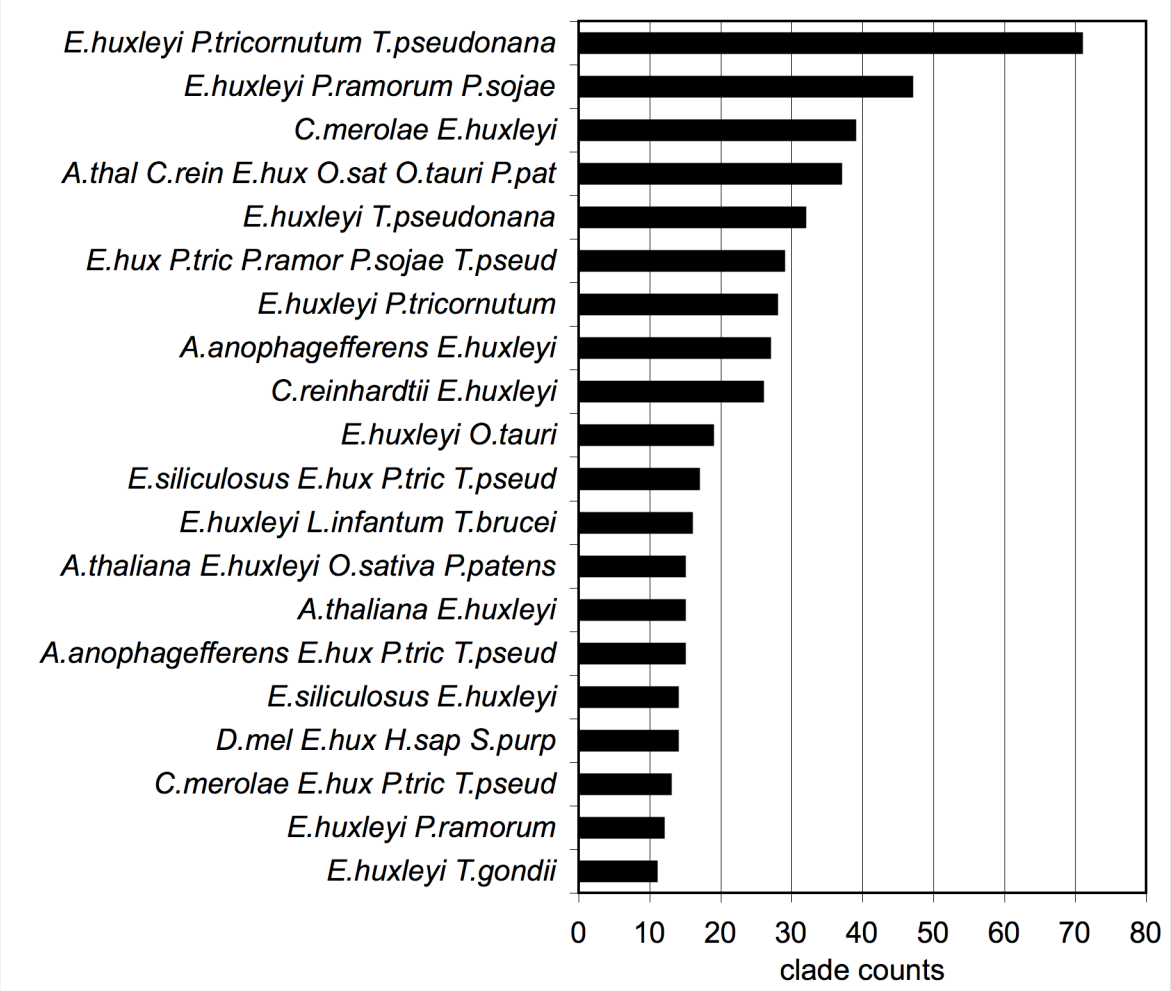


Figure 3: Counts of bipartitions with the exact contents shown on the left.

A.thal = *Arabidopsis thaliana*, *C.rein* = *Chlamydomonas reinhardtii*, *E.hux* = *Emiliania huxleyi*, *O.sat* = *Oryza sativa*, *P.pat* = *Physcomitrella patens*, *P.tric* = *Phaeodactylum tricornutum*, *P.ramor* = *Phytophthora ramorum*, *T.pseu* = *Thalassiosira pseudonana*, *D.mel* = *Drosophila melanogaster*, *H.sap* = *Homo sapiens*, *S.purp* = *Strongylocentrotus purpuratus*

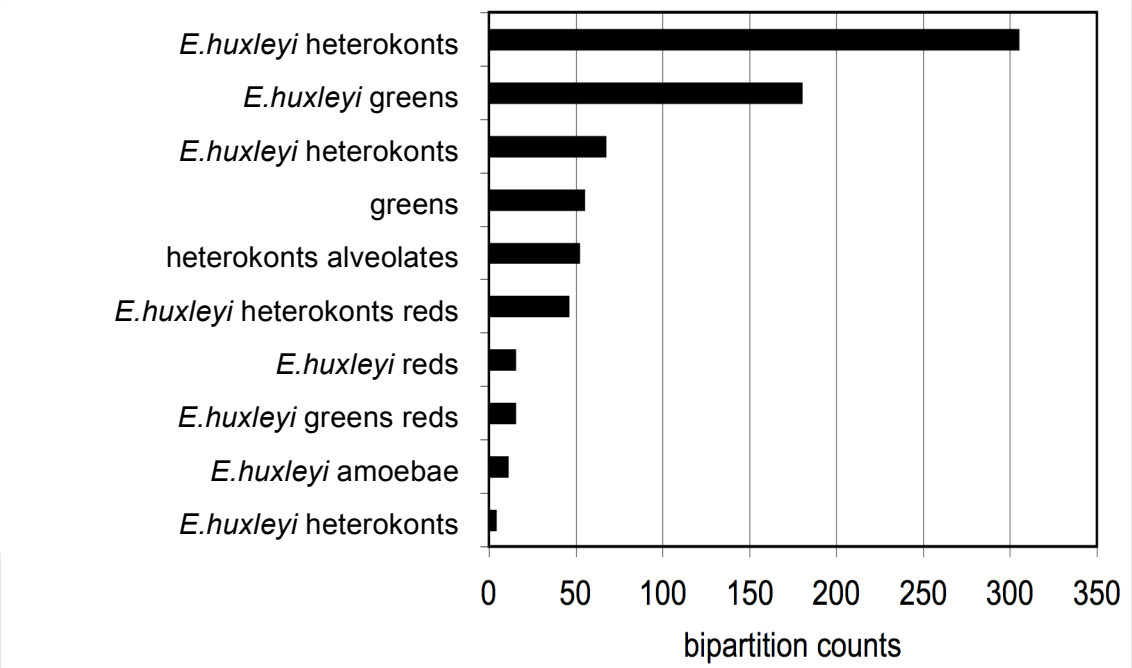


Figure 4: Counts of bipartitions including *E. huxleyi* and least one member of each of the other lineages.

To test for the possibility that that the variability in phylogenetic position of *E. huxleyi* was attributable to poor alignments, Needleman-Wunsch pairwise global alignments were done between each sequence in each COG and an *Emiliania* sequence in that COG. The distribution of Needleman-Wunsch scores is shown in figure 5. When sequences with global alignment scores below 800 were removed and new trees found the proportions of different bipartitions remained very similar although the numbers decreased (fig 6 and 7). To see if the taxa used could be narrowed based on their alignability to *Emiliania* sequences, organisms were counted based upon how often they were one of the top two aligned sequences to an *E. huxleyi* sequence in the COG. Quantification of the top two Needleman-Wunsch alignment scores in each COG to *E. huxleyi* by organism is shown in figure 8. As expected the most alignable sequences to *E. huxleyi* were from heterokonts and green/red lineage organisms. Surprisingly, animal sequences were among the most alignable sequences to *Emiliania* while alveolate sequences, which were expected to be alignable based on previous research, seemed less alignable to *E. huxleyi* sequences.

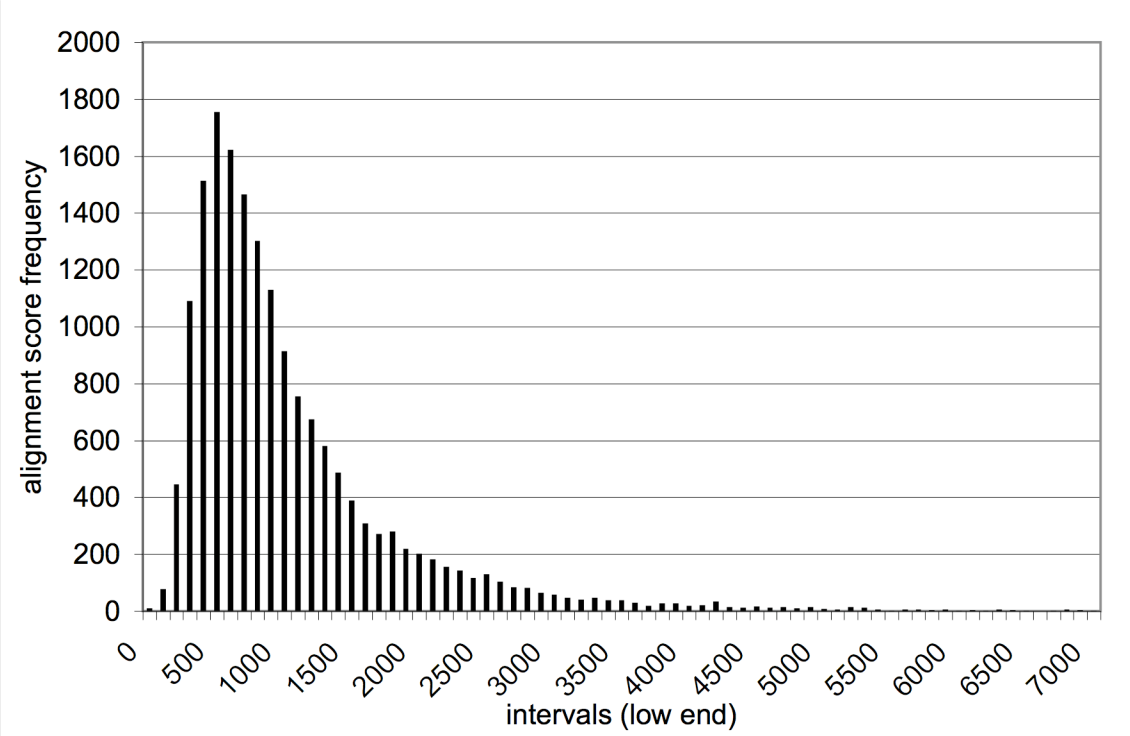


Figure 5: The distribution of Needleman-Wunsch global alignment scores of sequences in each COG to an *Emiliania huxleyi* sequence in that COG. Marks label low ends of intervals. The peak of 1756 sequence scores is between 500 and 600.

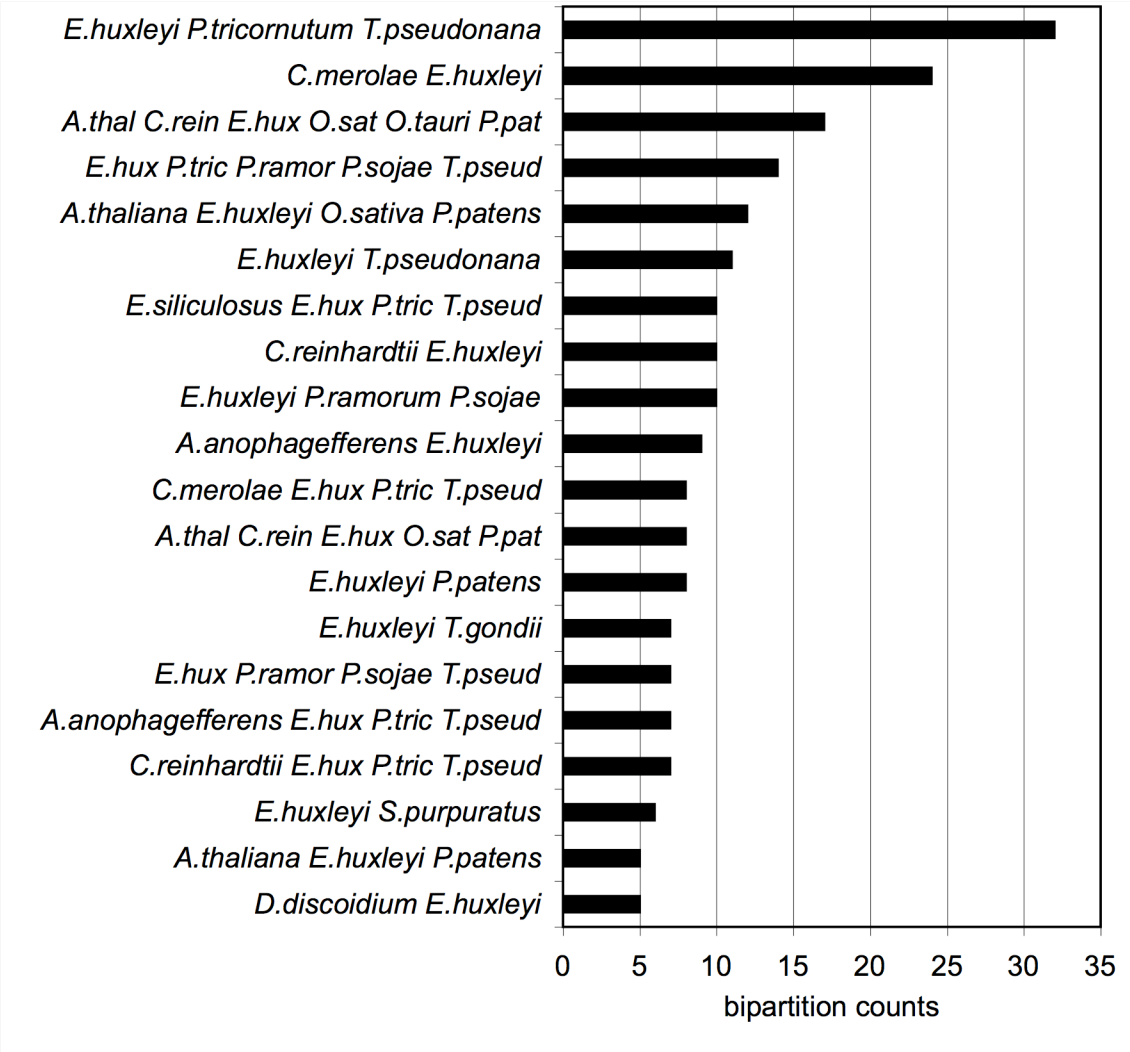


Figure 6: Counts of bipartition with the exact contents shown on the left from protein trees based on sequences with global alignment scores above 800 to an *E. huxleyi* sequence in each COG.

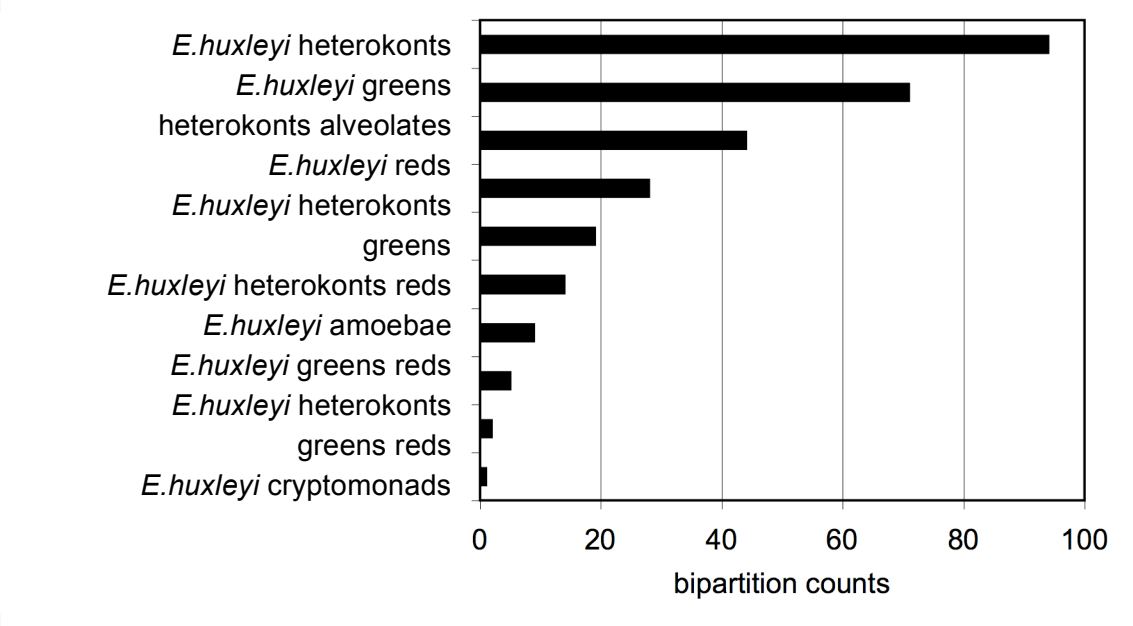


Figure 7: Counts of bipartitions including *E. huxleyi* and least one member of each of the other lineages on trees based on COGs from which sequences with Needleman-Wunsch scores below 800 had been removed.

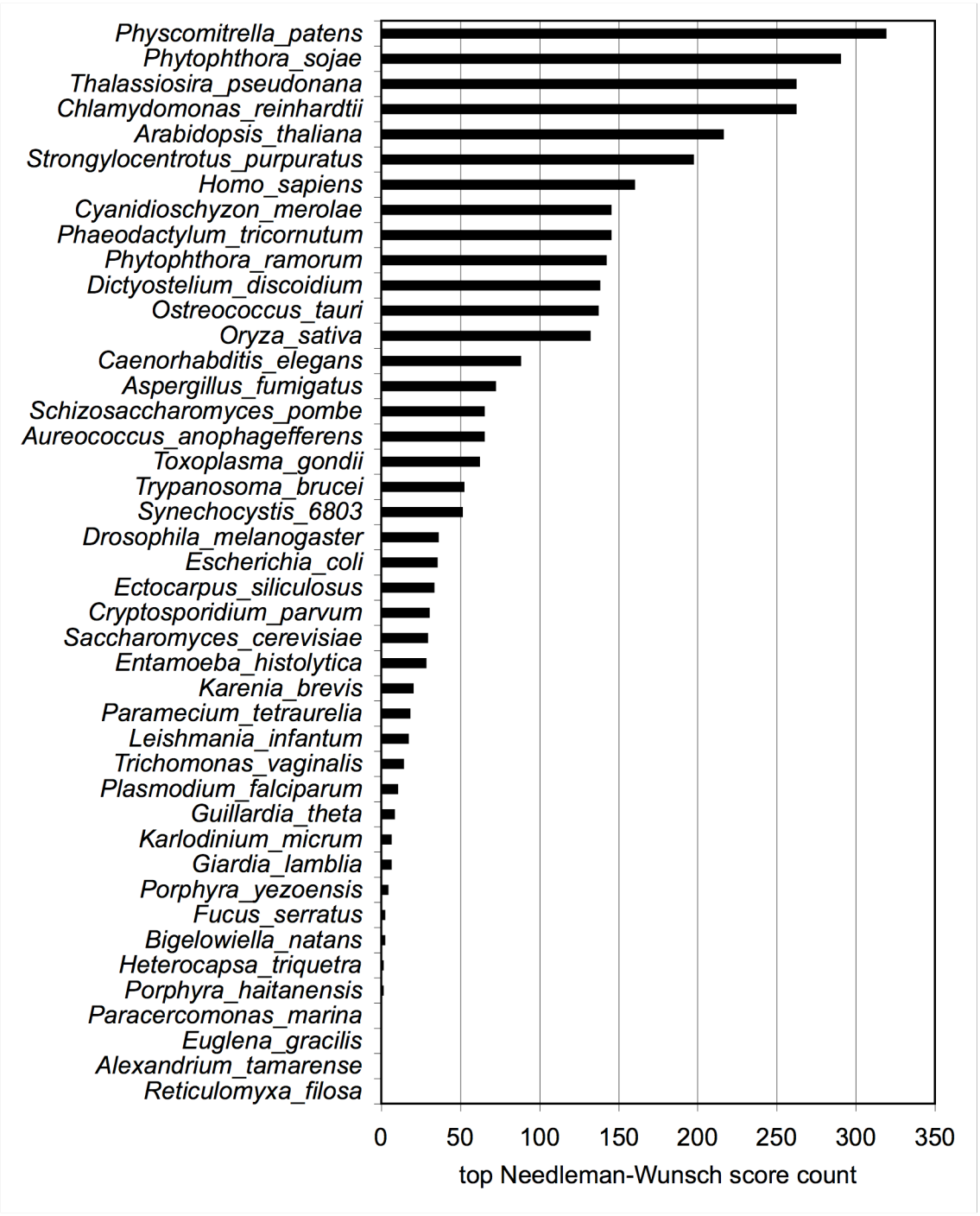


Figure 8: Quantification of the number of times sequences from each organism had one of the top two Needleman-Wunsch global alignment scores to the *E. huxleyi* sequence in a COG.

Emiliania huxleyi relationship with heterkonts

The 305 alignments that produced protein trees that put *E. huxleyi* in a bipartition with heterokonts to the exclusion of other taxa (for more information on alignments see table 2) were concatenated and paralogs removed. The resulting tree put *E. huxleyi* within the heterokonts as an outgroup to the phototrophic heterokonts, but with the oomycetes (*Phytophthora* spp.) as an outgroup to the phototrophic heterokonts plus *Emiliania* (fig 9). Bootstrap support within the heterokonts was high with the exception of the branch linking *Ectocarpus siliculosus* to the diatoms. Branches leading to and from *Emiliania* had 100% support. *Fucus serratus*, a heterokont, was located within the primary plastid containing bipartition where it was sister to *Porphyra haitanensis*, a rhodophyte. The cryptophyte, *Guillardia theta*, was an outgroup to the opisthokonts, not sister to *Emiliania*. Sequences with Needleman-Wunsch global alignment scores above 800 from each protein alignment were concatenated. This reduced data set was composed of 115 proteins (table 2). The phylogenetic tree resulting from this filtered concatenated alignment again put *E. huxleyi* within the heterokonts sister to the phototrophic heterokonts with high support (fig 10). When 20 random subsamples of 4000 amino acids were taken from the concatenated alignment, *Emiliania* was associated with the heterokonts in all resulting trees. *Guillardia theta* was associated with the heterokonts in only 12 of the trees and was only sister to *E. huxleyi* in one subsampled tree. Of the 305 genes linking *Emiliania* to the heterokonts 185 were represented in the *Arabidopsis thaliana*

genome. Within the 185 gene study set, mitochondrial and plastid genes were overrepresented compared to the population of genes which consisted of the genes in the *A. thaliana* genome (e-value < 0.05)(fig 11). The distribution of these genes on scaffolds was significantly different from the expected distribution based on scaffold length (Fisher's exact test: $p < 2.2e-16$) indicating that the proteins linking *E. huxleyi* to the heterokonts are physically clustered as genes in the genome of *E. huxleyi* (fig 12).

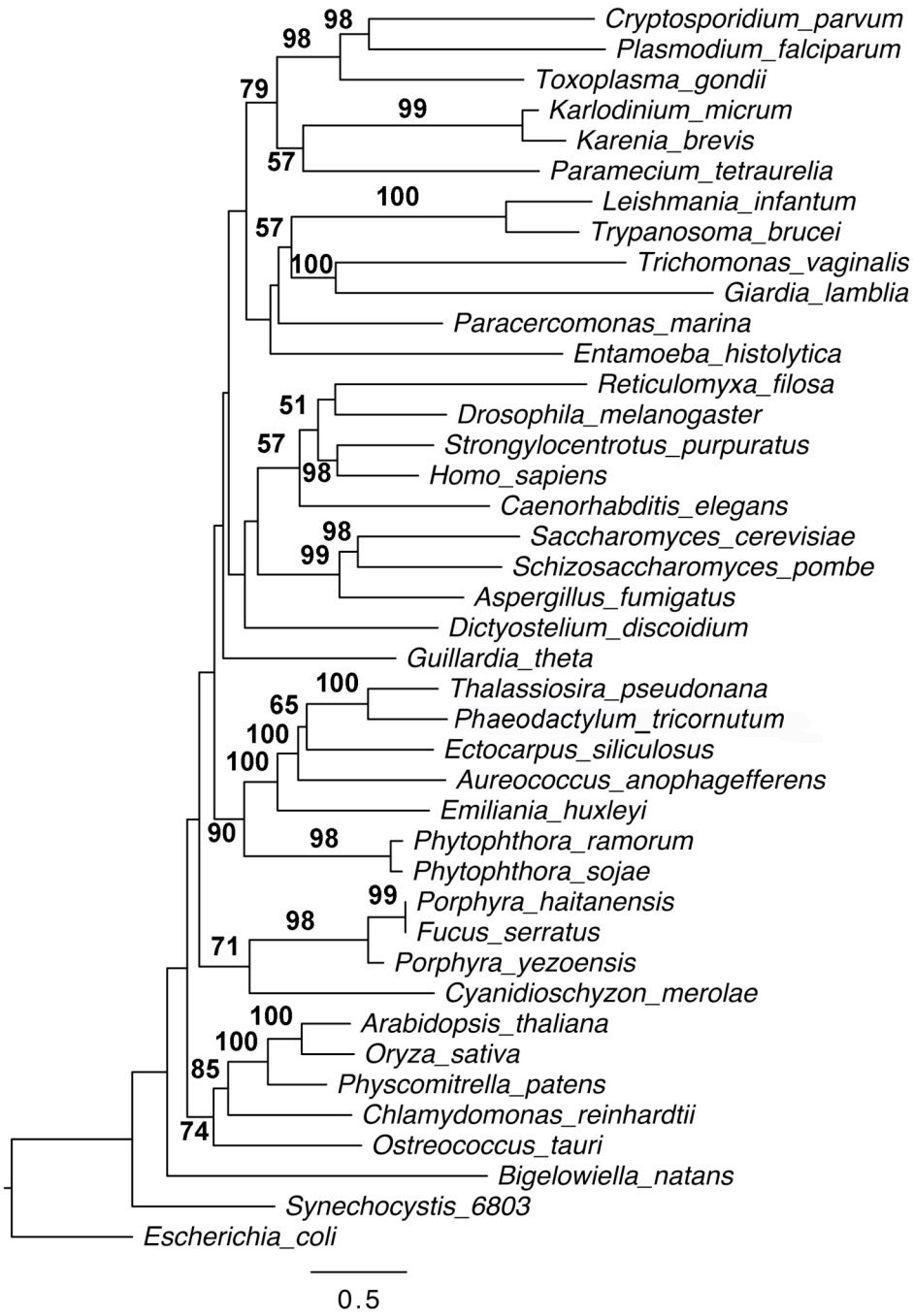


Figure 9: Phylogenetic likelihood tree resulting from analysis (JTT, Γ) of a 305 protein concatenated data set. Each individual protein tree grouped *E. huxleyi* with heterokonts. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 166 replicates.

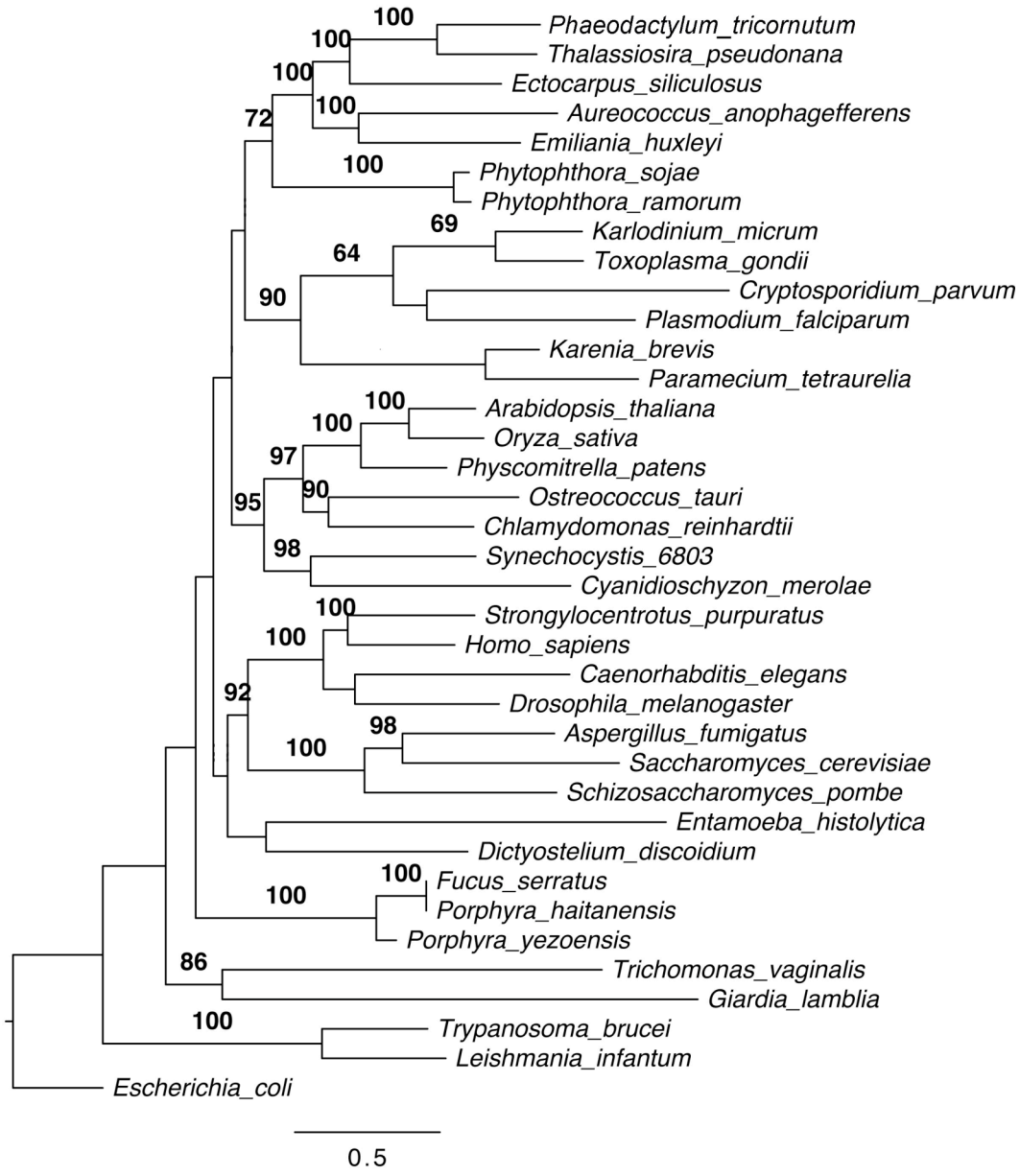


Figure 10: Phylogenetic likelihood tree resulting from analysis (JTT, Γ) of a 115 protein concatenated data set resulting from the removal of poorly aligned sequences from the 305 protein data set. Each individual protein tree supported the grouping of *E. huxleyi* with the heterokonts. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

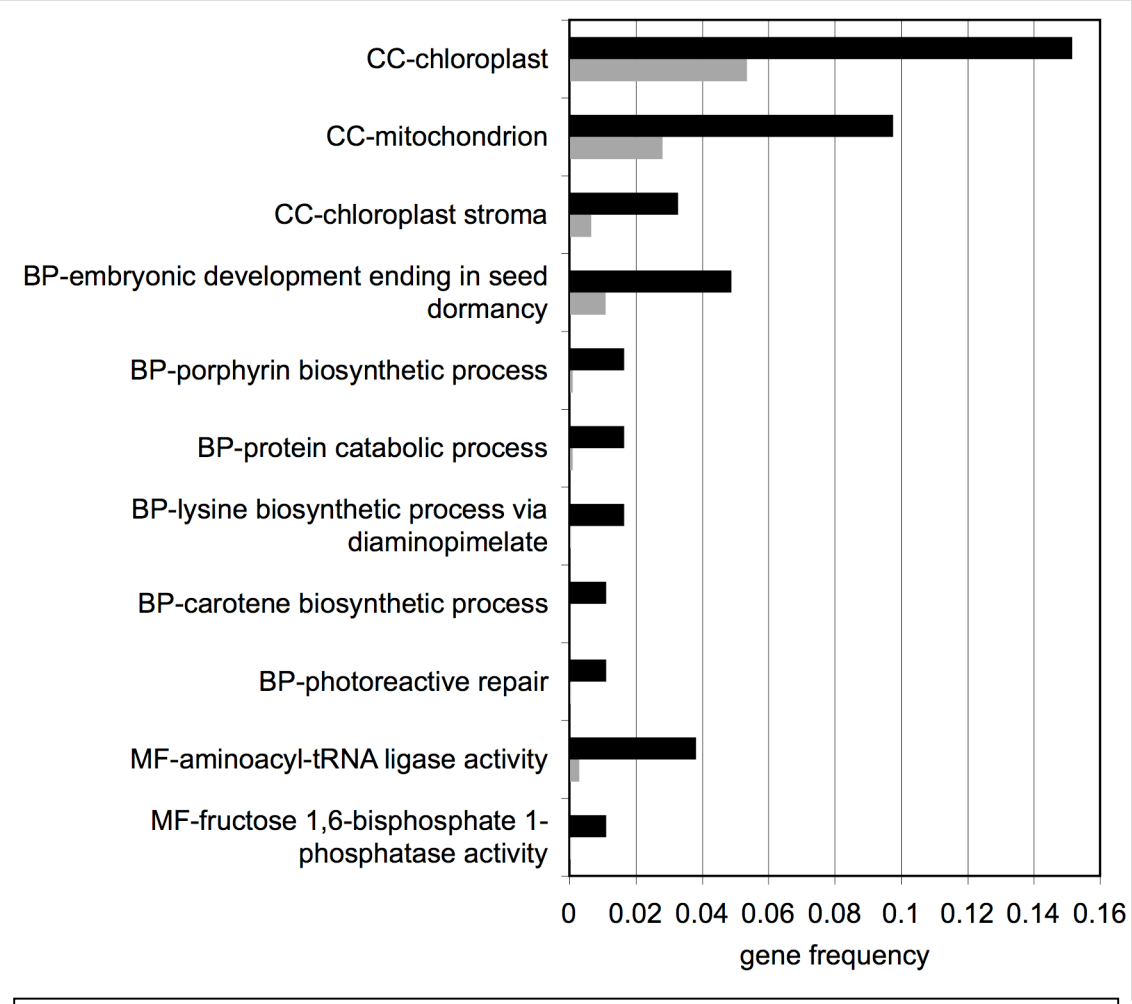


Figure 11: Gene merge results for genes linking *E. huxleyi* to heterokonts. Gene Ontology (GO) categories in which study genes (black bars) were overrepresented (e-value < 0.05) compared to gene population (genes of *A. thaliana* genome)(grey bars). The study genes were composed of the 185 genes in the *A. thaliana* genome that were associated with the 305 genes linking *Emiliania* to the heterokonts. CC = cell component 161 genes with GO terms, BP = biological process 166 genes with GO terms, MF = molecular function 172 genes with GO terms.

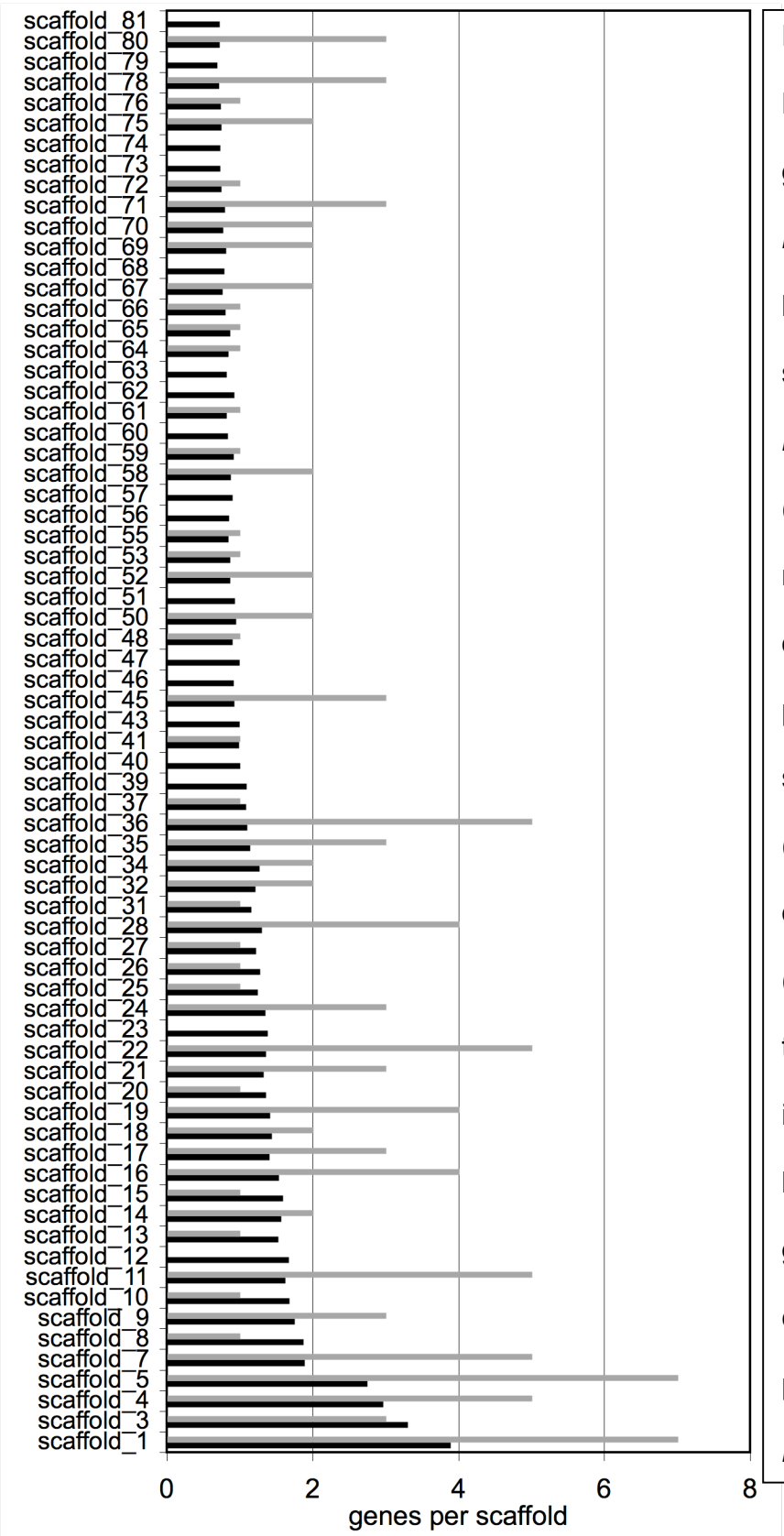


Figure 12: Distribution of genes linking *E. huxleyi* to heterokonts on scaffolds of *E. huxleyi* genome (grey bars) and the number of genes on each scaffold predicted by scaffold length (black bars). The difference (Fisher's exact test: $p < 2.2e-16$) implies that the heterokont related genes are clustered physically in the *E. huxleyi* genome.

Emiliania huxleyi relationship with Green lineange

The 180 alignments that grouped *Emiliania* with members of the Viridiplantae (green lineage) were concatenated and the paralogs were removed (see table 2 for additional information about alignments). The phylogenetic analysis put *E. huxleyi* as an outgroup to the Viridiplantae with high support (fig 13). Sister to the green-*E. huxleyi* bipartition were the rhodophytes (red lineage) and the cyanobacterium *Synechocystis* but without significant support. *Guillardia theta* grouped with the heterokonts but without significant support. When sequences with Needleman-Wunsch scores below 800 were removed, 87 protein alignments remained. These 87 were concatenated and a best tree found. The resulting tree again put *E. huxleyi* directly outside the Viridiplantae with high support (fig 14). Trees based on 4000 amino acid subsamples from the concatenated alignment resulted in *E. huxleyi* grouping with the greens nine times, the heterokonts six times, and the alveolates seven times. The heterokonts associated with the greens seven times and the alveolates five times. Gene merge results indicate that genes linking *E. huxleyi* to the green lineage were enriched in a few GO categories (e-value < 0.05)(fig 15). Genes associating *E. huxleyi* with the Viridiplantae are clustered disproportionately on some scaffolds implying possible synteny (Fisher's exact test: $p < 2.2e-16$)(fig 16).

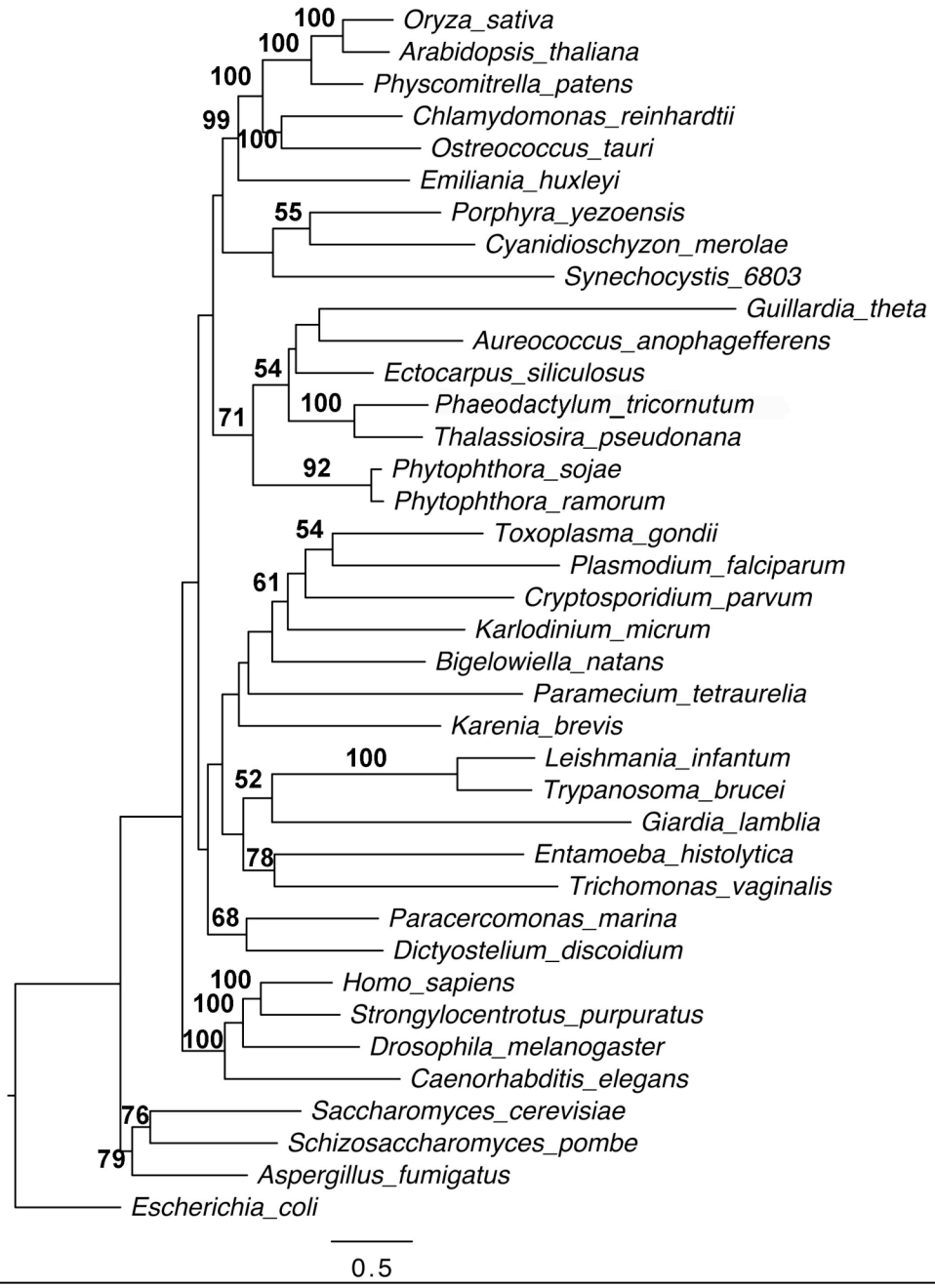


Figure 13: Likelihood tree resulting from analysis (JTT, Γ) of a 180 protein concatenated data set. Each individual protein tree grouped *E. huxleyi* with green lineage organisms. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 192 replicates.

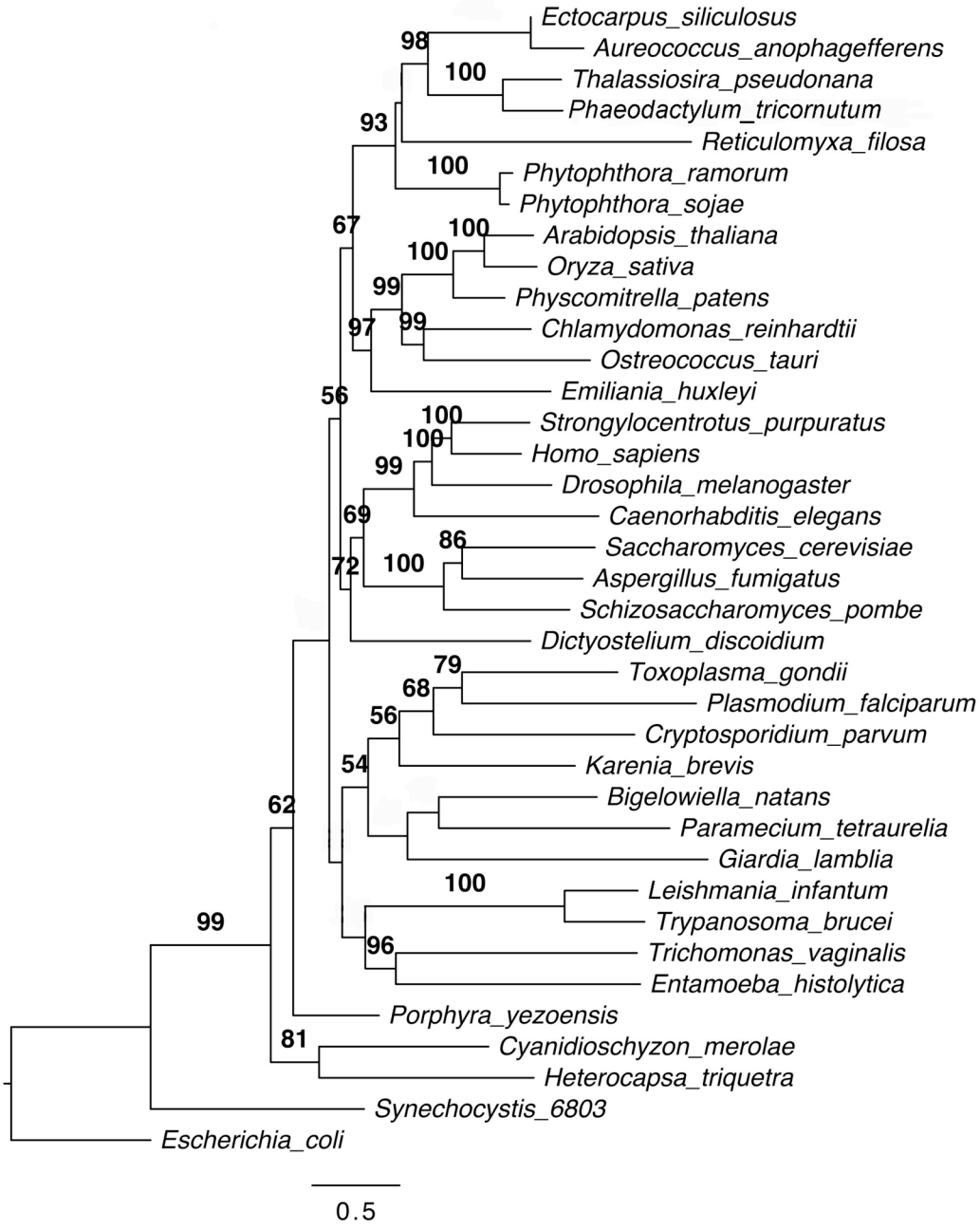


Figure 14: Likelihood tree resulting from analysis (JTT, Γ) of an 87 protein concatenated data set resulting from the removal of poorly aligned sequences from the 180 protein data set. Each individual protein tree supported the grouping of *E. huxleyi* with the Viridiplantae. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

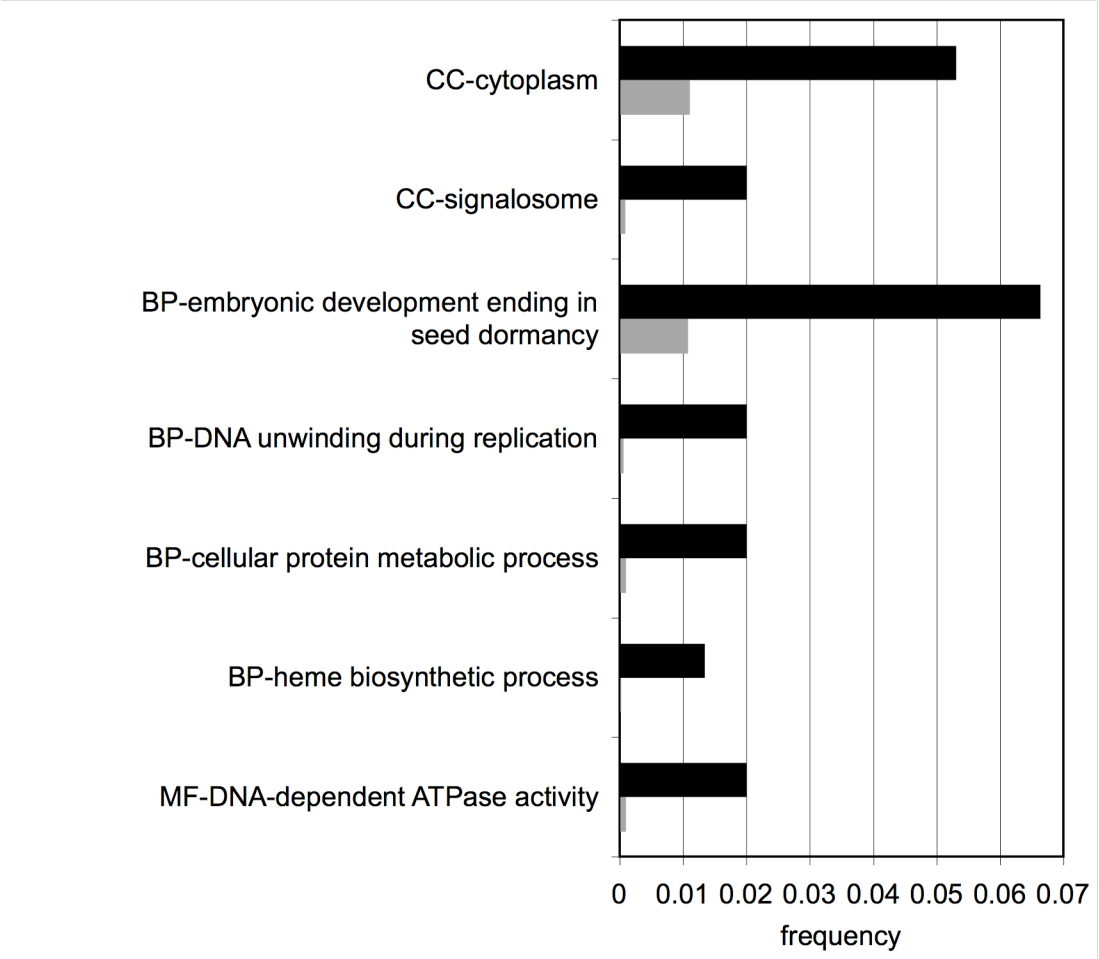


Figure 15: Gene merge results for genes linking *E. huxleyi* to the Viridiplantae. Gene Ontology (GO) categories in which study genes (black bars) were overrepresented (e-value < 0.05) compared to gene population (genes of *A. thaliana* genome)(grey bars). The study genes were composed of the 151 genes in the *A. thaliana* genome that were associated with the 180 genes linking *Emiliania* to the Viridiplantae. CC = cell component 124 genes with GO terms, BP = biological process 143 genes with GO terms, MF = molecular function 129 genes with GO terms.

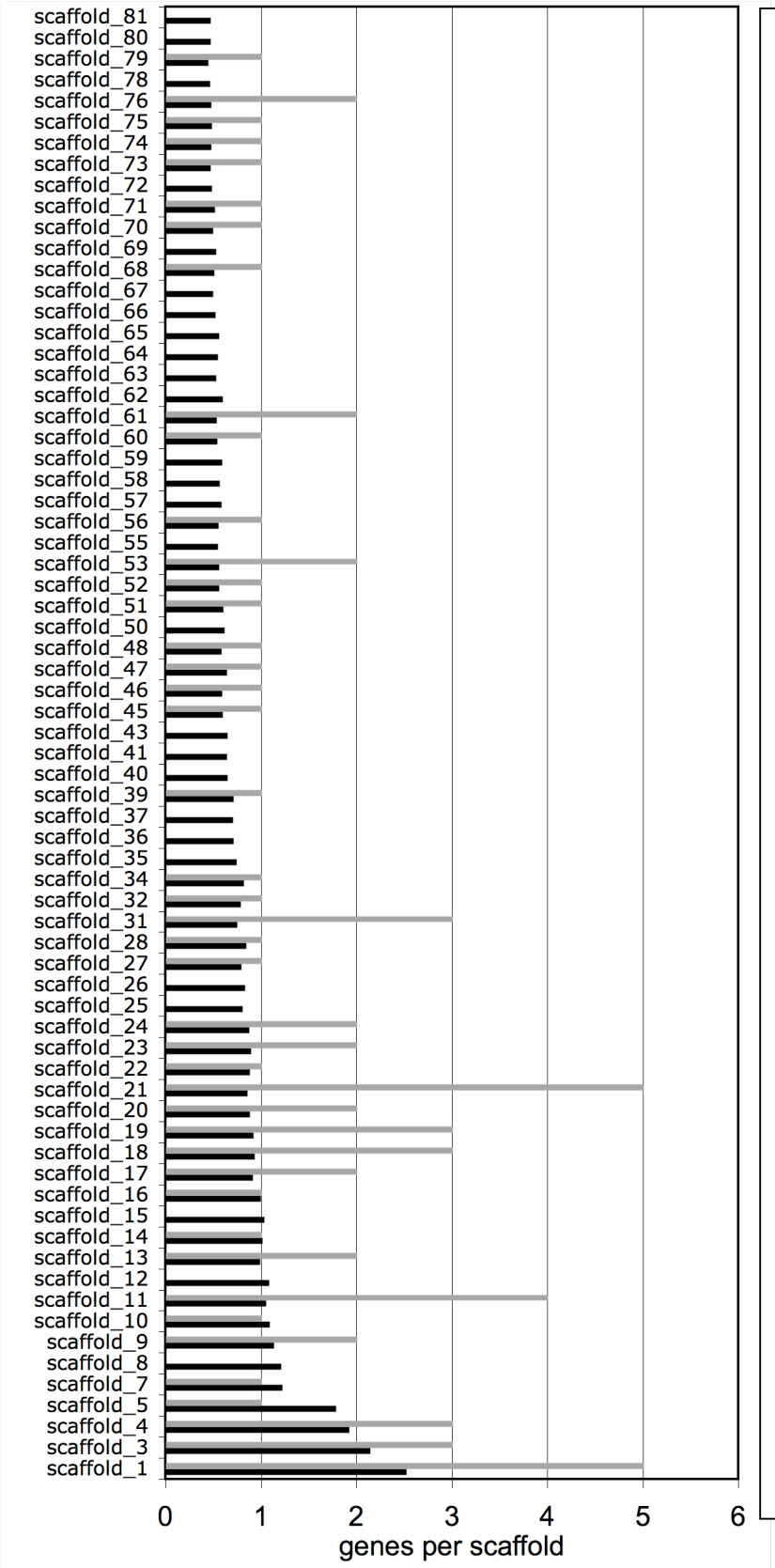


Figure 16: Distribution of genes linking *E. huxleyi* to green lineage organisms on scaffolds of *E. huxleyi* genome (grey bars) and the number of genes on each scaffold predicted by scaffold length (black bars). The difference (Fisher's exact test: $p < 2.2e-16$) implies that the Viridiplantae related genes are clustered physically in the *E. huxleyi* genome.

Emiliania huxleyi relationship with Reds

The concatenated alignment included 46 protein alignments that individually resulted in protein trees grouping *Emiliania* with the rhodophytes (See table 2 for additional information on alignments.). The resulting likelihood tree put *E. huxleyi* and *Karlodinium micrum* as sister taxa with bootstrap support slightly below the 70 that is normally considered significant. This pair in turn was sister to the rhodophytes *Porphyra yezoensis* and *Cyanidioschyzon merolae* again with mediocre bootstrap support (fig 17). The filtered data produced a concatenated alignment that included 29 separate protein alignments. This tree shows *E. huxleyi* as sister to *P. yezoensis* with *C. merolae* as the outermost red algae all with poor support (fig 18). *K. micrum* was included within the alveolates in the filtered tree. The only two GO categories that Gene Merge identified as being over represented in this group of genes were the cell component categories mitochondrion and nucleolus (e-value < 0.05). The test for clustering of rhodophyte associated genes on scaffolds was not significant, possibly due to the smaller number of genes in this category.

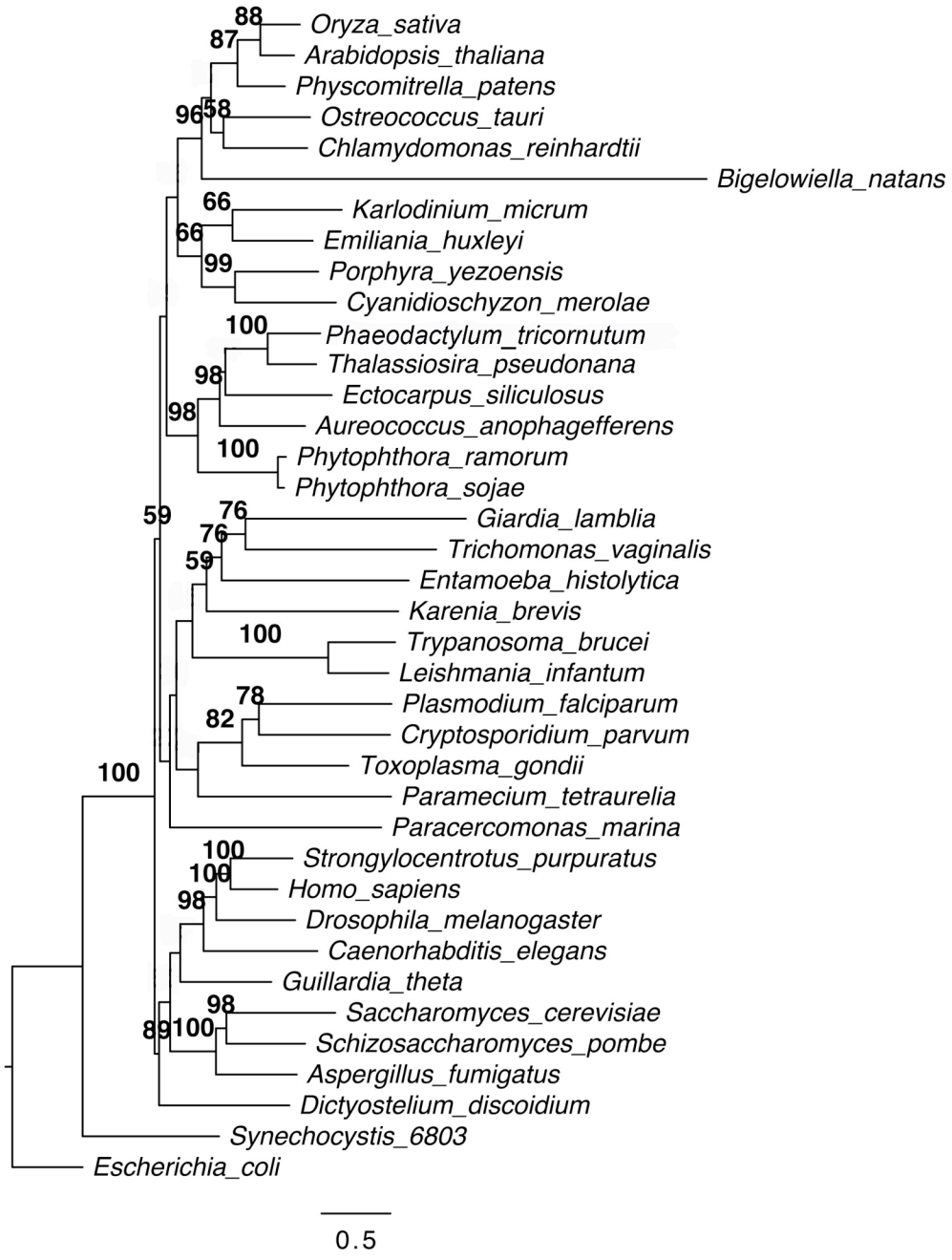


Figure 17: Likelihood tree resulting from analysis (JTT, Γ) of a 46 protein concatenated data set. Each individual protein tree grouped *E. huxleyi* with red lineage organisms. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

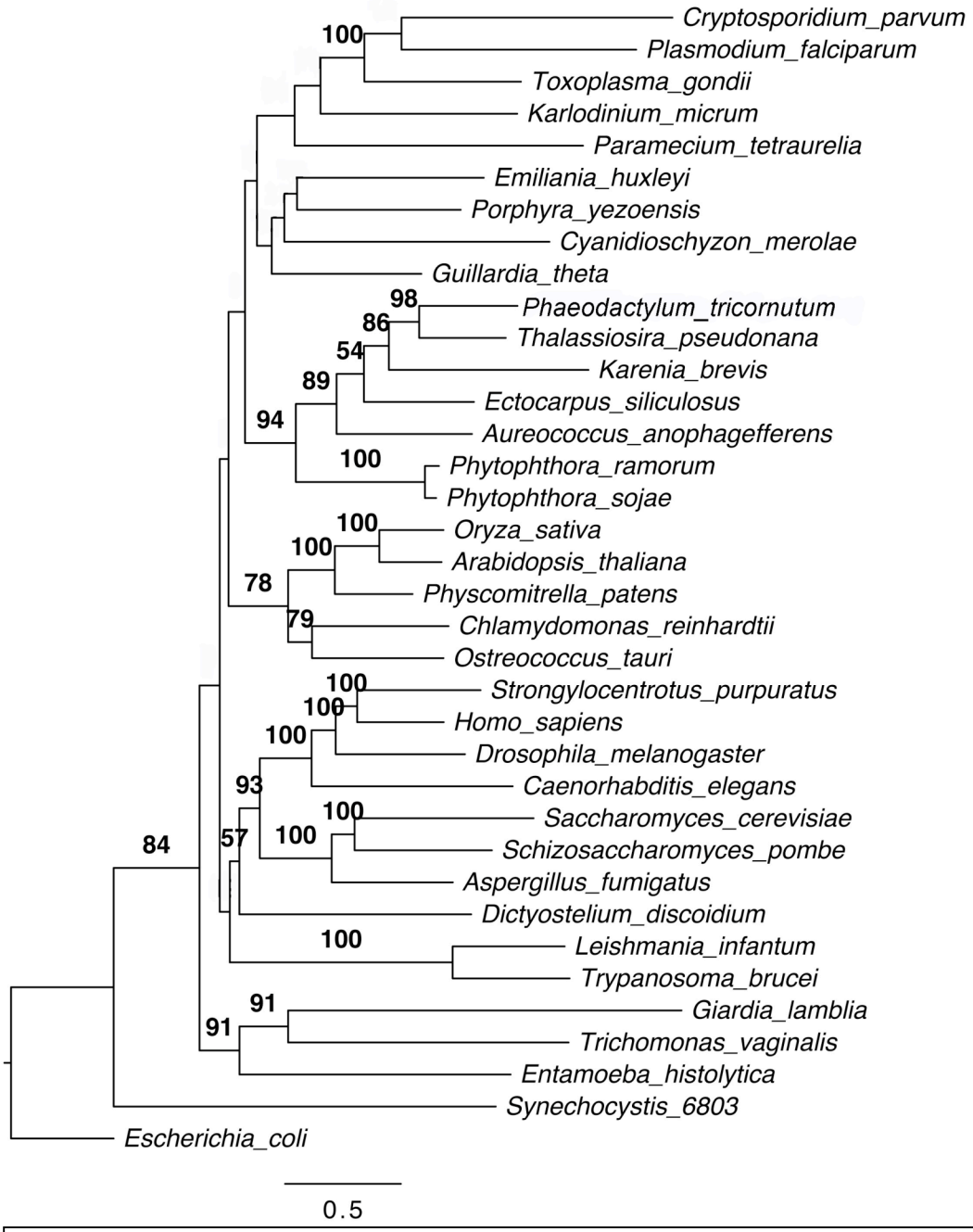


Figure 18: Likelihood tree resulting from analysis (JTT, Γ) of a 29 protein concatenated data set after the removal of poorly aligned sequences from the 46 protein data set. Each individual protein tree supported the grouping of *E. huxleyi* with the Rhodophytes. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

Emiliania huxleyi and cryptophytes

There were four proteins that individually supported a sister relationship between *Emiliania huxleyi* and *Guillardia theta* (For additional information on alignments see table 2.). The concatenated alignment produced a likelihood tree that strongly supported a relationship between *E. huxleyi* and the cryptophyte *Guillardia theta* (fig 19). After sequences with Needleman-Wunsch alignment scores below 800 were removed, there was only one *G. theta* gene left and support for the *E. huxleyi* sister relationship with *G. theta* was still barely significant (fig 20). It is worth remembering that *G. theta* is represented by a relatively small number of sequences in this study.

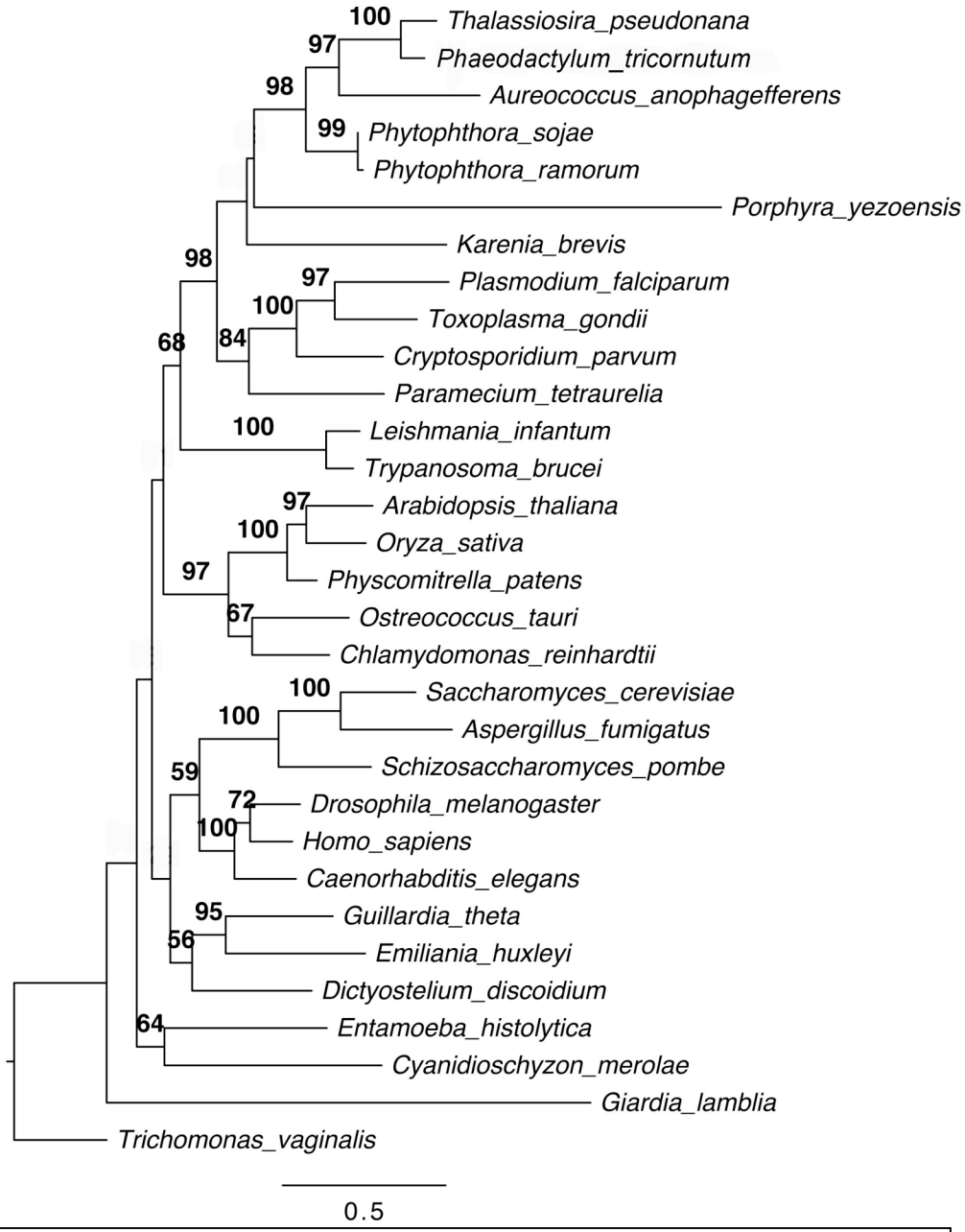


Figure 19: Likelihood tree resulting from analysis (JTT, Γ) of a four protein concatenated data set. Each individual protein tree grouped *E. huxleyi* with the cryptophyte *Guillardia theta*. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

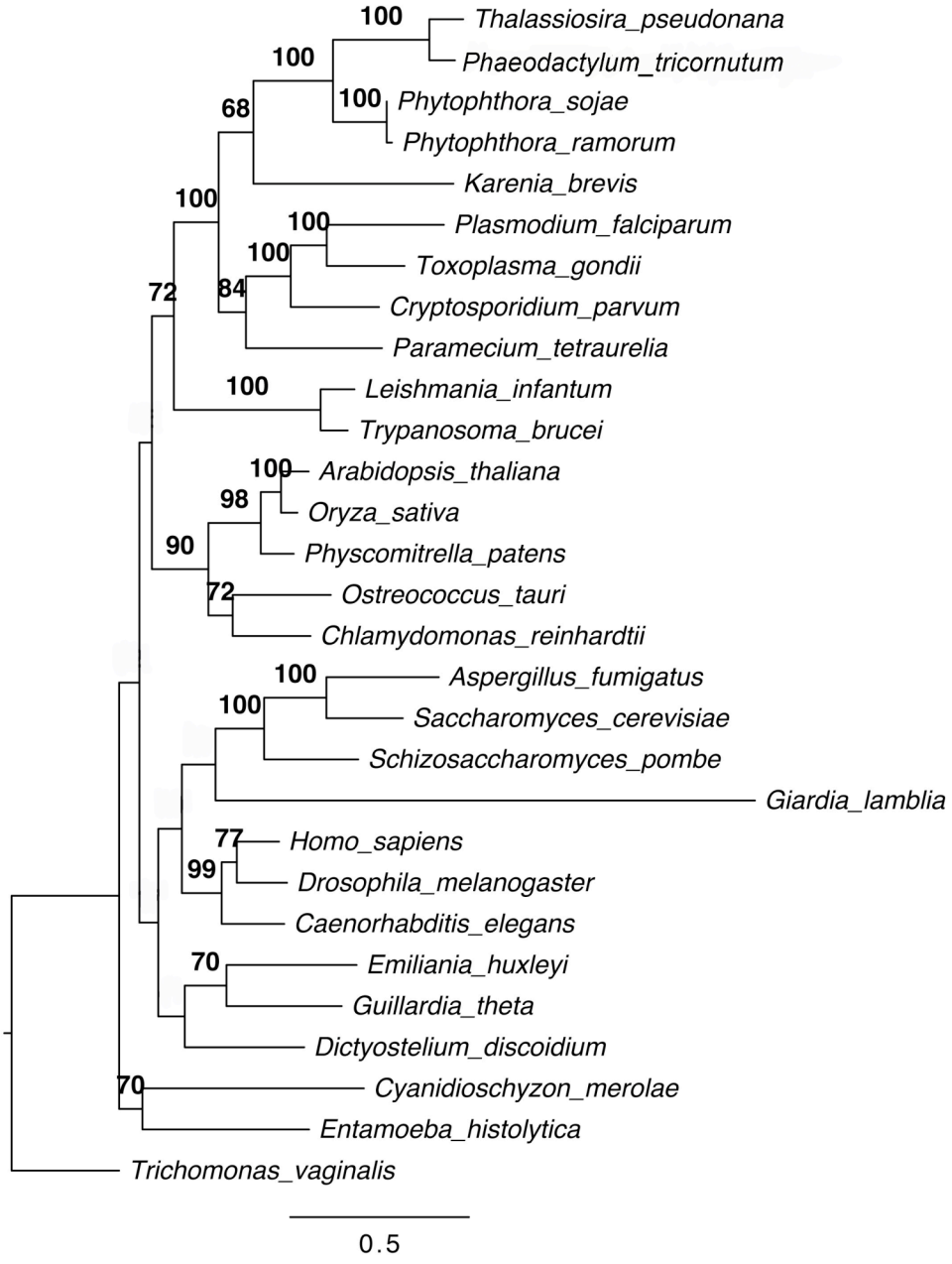


Figure 20: Likelihood tree resulting from analysis (JTT, Γ) of a single protein data set after the removal of poorly aligned sequences from the four protein data set. Each individual protein tree supported the grouping of *E. huxleyi* with the cryptophyte *G. theta*. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

Alveolates with heterokonts

A relationship with good support was found between heterokonts and alveolates based on a concatenated alignment of 55 proteins (fig 21)(For additional information on alignments see table 2.). *Bigelowiella natans* was sister to *Paramecium tetraurelia*, though with poor support, and together they formed an outgroup to the dinoflagellates and apicomplexans. This alignment also has *E. huxleyi* as sister to *Guillardia theta*, consistent with previous studies. When sequences with Needleman-Wunsch alignment scores below 800 compared to an *E. huxleyi* sequence in the same COG were removed, the resulting tree still provided good support for a sister relationship between the alveolates and the heterokonts with *B. natans* still sister to *P. tetraurelia* (fig 22). *E. huxleyi* ended up sister to *E. coli* in this tree.

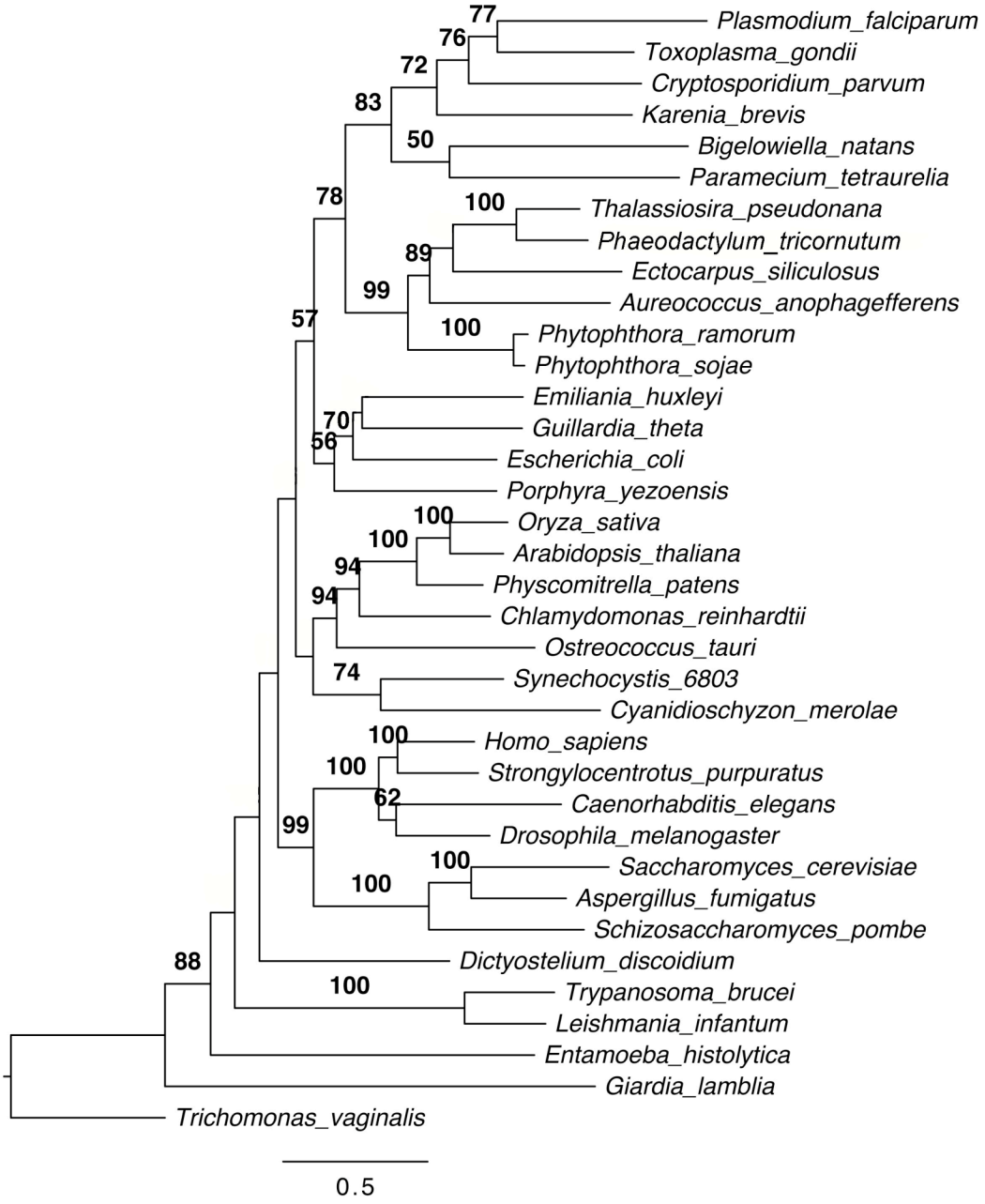


Figure 21: Likelihood tree resulting from analysis (JTT, Γ) of a 55 protein concatenated data set. Each individual protein tree grouped alveolates with heterokonts. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

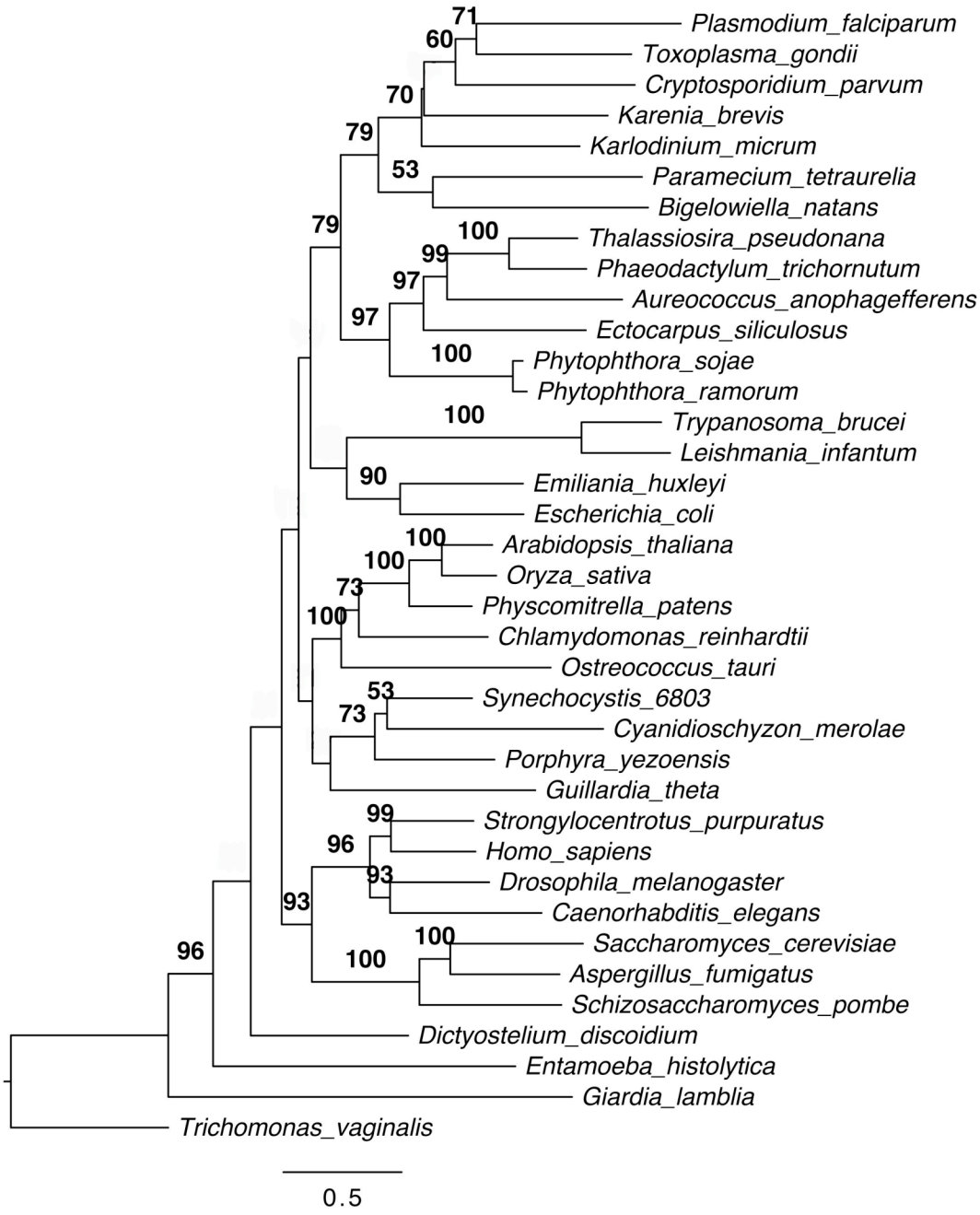


Figure 22: Likelihood tree resulting from analysis (JTT, Γ) of a 44 protein data set after the removal of poorly aligned sequences from the 55 protein data set. Each individual protein tree supported the grouping of alveolates with heterokonts. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

Emiliania with the greens and heterokonts

There were 67 individual protein trees that put *E. huxleyi* in a bipartition with the heterokonts and the Viridiplantae. When the alignments were concatenated (For additional information on alignments see table 2.) and analyzed together, *E. huxleyi* grouped with the heterokonts with high support and the rhodophytes were sister to the heterokonts (fig 23). The green lineage was sister to the *Emiliania*-heterokont-rhodophyte bipartition but without strong support. The COGs were then filtered so the sequences with Needleman-Wunsch scores below 800 to *Emiliania* were removed. When the alignment and analysis was run again with the 19 protein filtered dataset, *Emiliania* again fell out with the heterokonts with good support and the alliance with the green lineage lacked good support (fig 24).

Emiliania with the reds and heterokonts

Emiliania huxleyi grouped with the heterokonts and the rhodophytes in 52 separate protein trees. These alignments were concatenated (For additional information on alignments see table 2.) and analyzed and the resulting tree grouped *E. huxleyi* within the heterokonts with high support (fig 25). *Cyanidioschyzon* was sister to the heterokont-*E. huxleyi* bipartition with moderately good support. When sequences with Needleman-Wunsch scores below 800 to *Emiliania* were filtered out, there were 14 proteins remaining, and the resulting tree again put *E. huxleyi* within the heterokonts and put *Porphyra yezoensis* sister to *Phytophthora* spp. (fig 26).

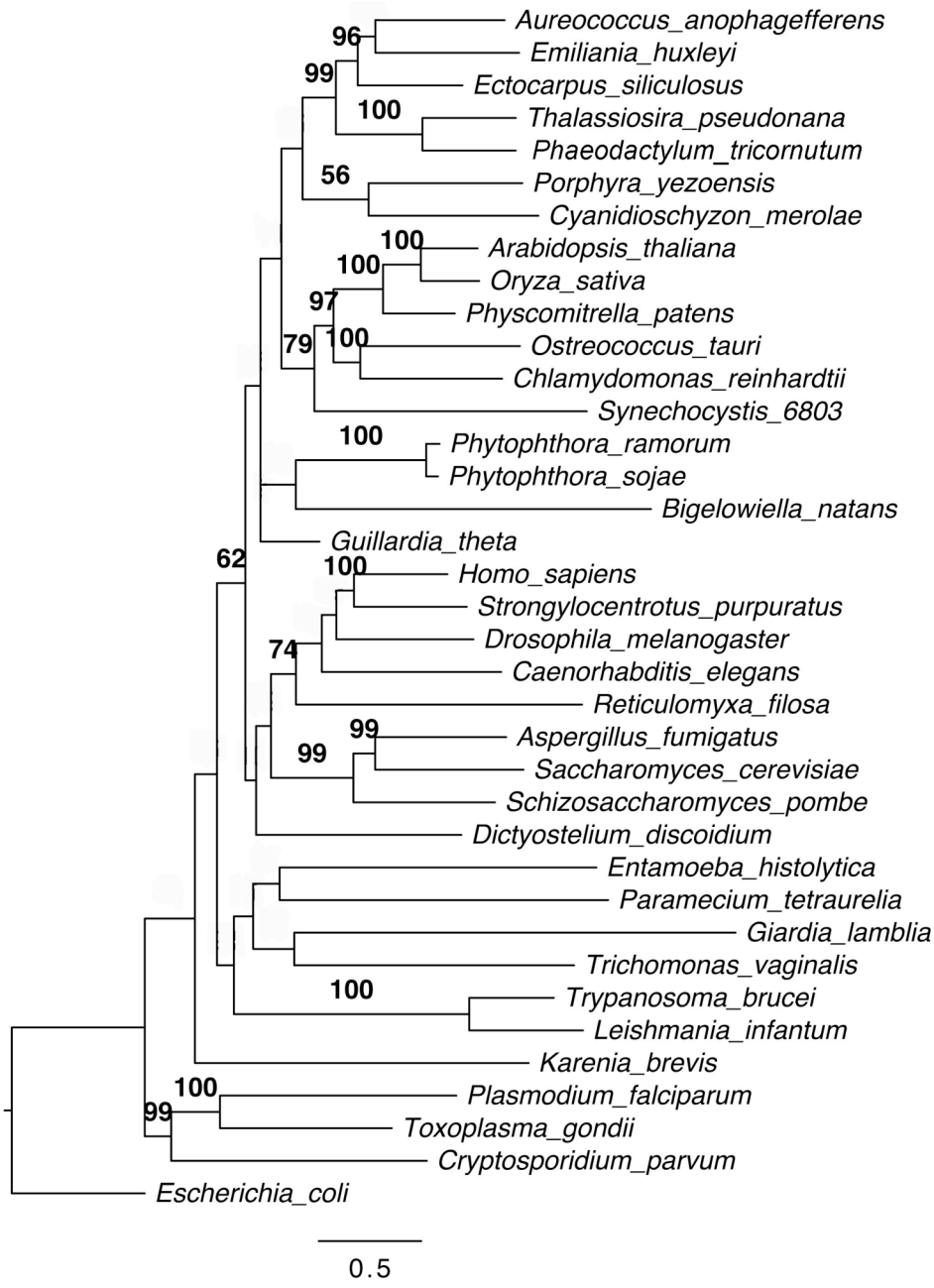


Figure 23: Likelihood tree resulting from analysis (JTT, Γ) of a 67 protein concatenated data set. Each individual protein tree grouped *E. huxleyi* with heterokonts and Viridiplantae. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

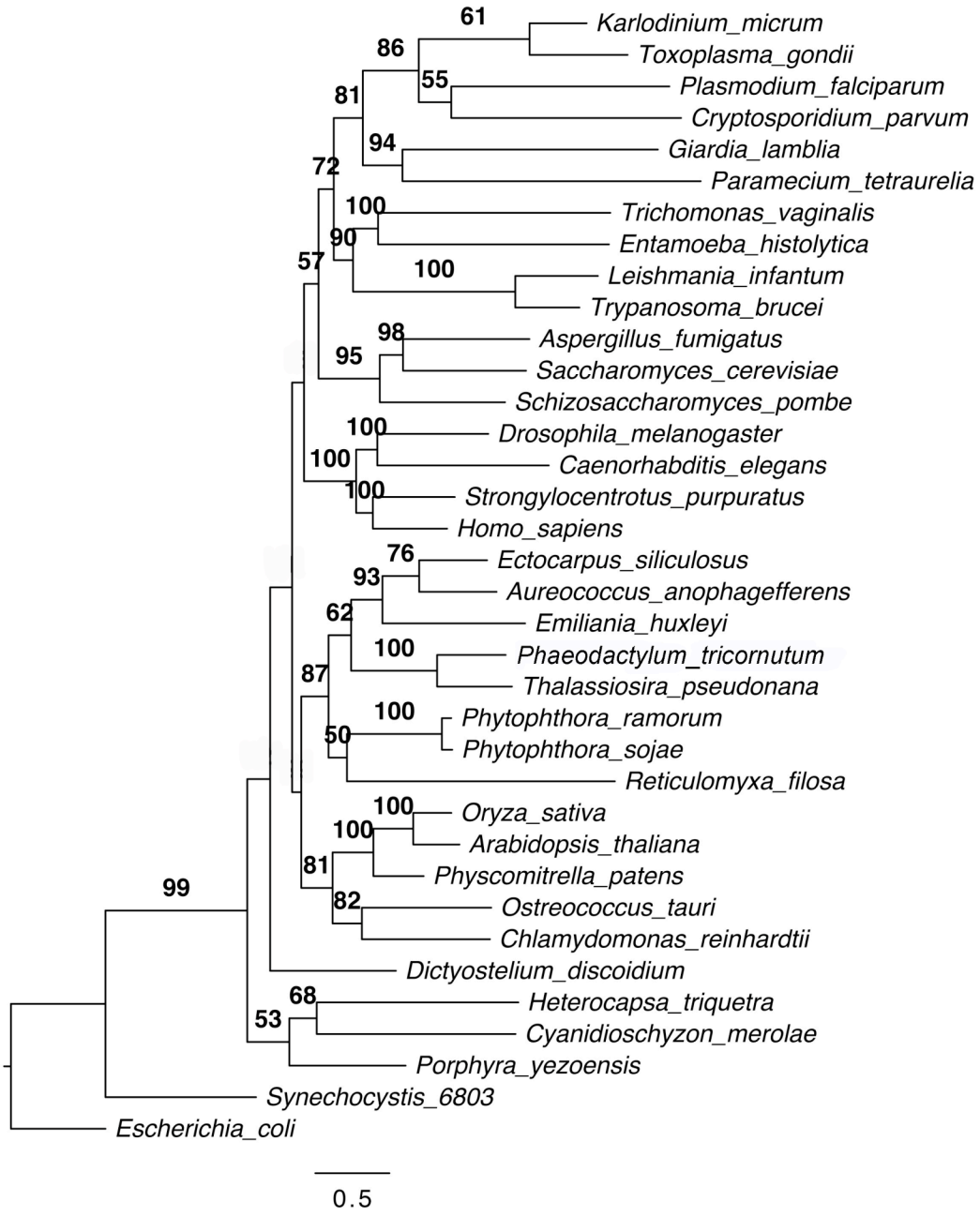


Figure 22: Likelihood tree resulting from analysis (JTT, Γ) of a 19 protein data set after the removal of poorly aligned sequences from the 67 protein data set. Each individual protein tree supported *Emiliania* grouping with heterokonts and Viridiplantae. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

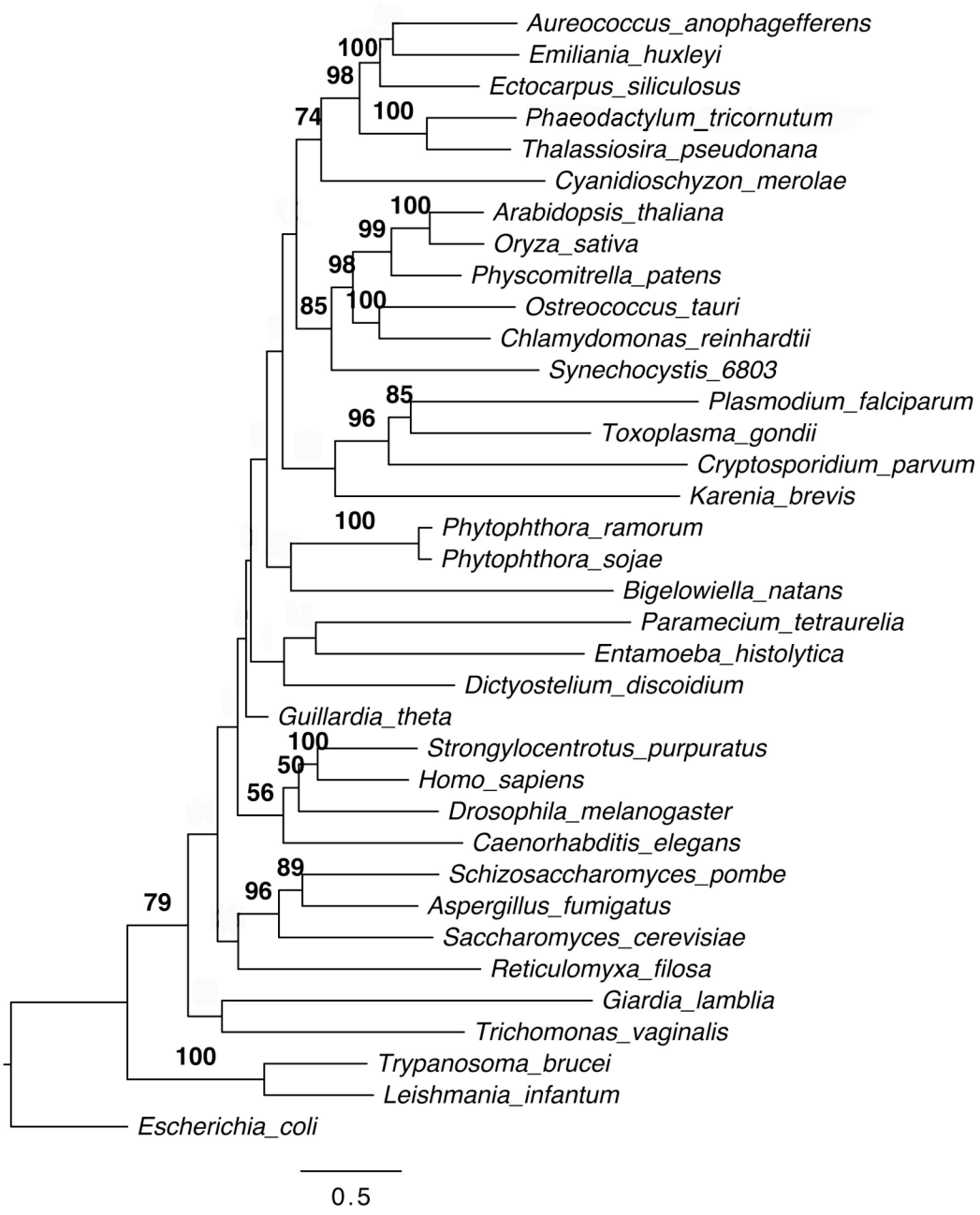


Figure 25: Likelihood tree resulting from analysis (JTT, Γ) of a 52 protein concatenated data set. Each individual protein tree grouped *E. huxleyi* with heterokonts and rhodophytes. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

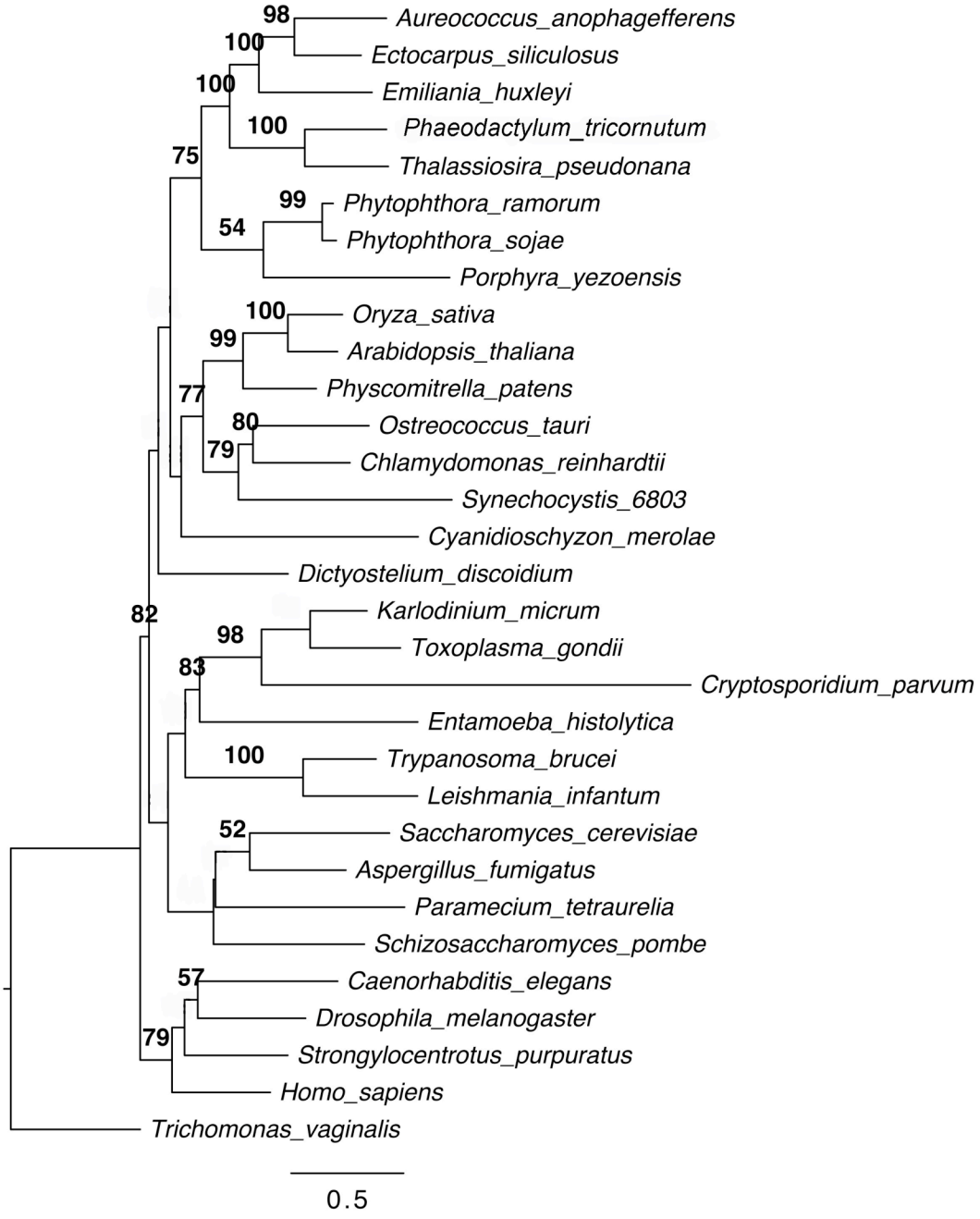


Figure 26: Likelihood tree resulting from analysis (JTT, Γ) of a 14 protein data set after the removal of poorly aligned sequences from the 52 protein data set. Each individual protein tree supported an *E. huxleyi*-heterokont-rhodophyte bipartition. Scale bar represents RAxML evolutionary distance. Bootstrap values based on 105 replicates.

Discussion

This study resulted in several large data sets that strongly support the affiliation of *Emiliania huxleyi* with several major groups of organisms. The chimeric nature of the *E. huxleyi* genome means that more than one of these phylogenetic placements could simultaneously be correct, but it is also important to consider potential sources of error, including phylogenetic artifacts, hidden paralogy, gene transfer, and the limitations of genome sampling. The largest dataset put *E. huxleyi* in a bipartition with the heterokonts. There was also strong support for grouping this haptophyte with green algae, with red algae, or with cryptomonads (=cryptophytes). Physical clustering of genes supporting particular bipartitions on some scaffolds implies synteny and bolsters the argument that *E. huxleyi* is chimeric as does the limited number of GO terms showing significant enrichment in genes from the datasets grouping *E. huxleyi* separately with the heterokonts, green algae, and red algae.

Discrepancies between trees

There are several types of systematic error that could result in disagreement among gene trees. Examples are variation in tree building method, nucleotide composition bias, and saturation (multiple changes at individual positions in a sequence). Method variation was not a problem in this case because in any given analysis the same method was used. Another methodological problem that might produce anomalous bipartitions is the use of poor alignments. For this reason, each sequence in each COG was aligned to

an *Emiliania* sequence in that COG and sequences with Needleman-Wunsch scores below 800 were removed and the analysis repeated on the filtered data. The Needleman-Wunsch algorithm (Needleman and Wunsch 1970) was used because it produces global alignment scores. Finding phylogenetic trees from multiple sequence alignments assumes that the sequences represent the same gene and they should be alignable from end to end. Removing sequences based on poor global alignment scores makes the same assumption and seemed to be the more conservative approach compared to using a local alignment procedure such as Smith-Waterman (Smith and Waterman 1981). In both cases the proportions of the most common associations with *E. huxleyi* were similar although the total counts were smaller. One curious result of the filtration of the alignments linking *E. huxleyi* to the red algae is that both concatenated trees grouped *Emiliania* with the rhodophytes, but that the filtered alignment resulted in a tree with lower bootstrap support. This implies that poor alignments do not always result in reduced support. The filtration of this data also moved *Karlodinium micrum* from the red lineage to the alveolates. Both of these placements are reasonable because the dinoflagellate *K. micrum* is an alveolate, but its plastid is of rhodophyte and more recently haptophyte origin (Tengs *et al*. 2000). The effect of composition bias and saturation was reduced by the used of amino acid sequences rather than nucleotides and by the use of likelihood rather than parsimony analytical methods. Stochastic error can also be a problem but is usually reduced by the use of large amounts of data (Jeffroy *et al*. 2006) and would produce random associations, whereas this pipeline puts *Emiliania* in a

few bipartitions with high frequency. I believe these high frequency associations are reliable while the lower frequency bipartitions should be treated more carefully.

Three of the remaining reasons for discrepancies between gene trees and species trees are: 1) paralogy, 2) lineage sorting/deep coalescence, and 3) horizontal gene transfer (Delwiche and Palmer, 1996; Maddison 1997). Paralogy by itself should, in most cases, simply increase counts of bipartitions that occur anyway, not produce aberrant bipartitions. The pipeline used in this study counted overlapping bipartitions and did not require all members to be present and so gene loss should not affect the results. Lineage sorting could produce aberrant bipartitions, but lineage sorting is most prevalent when diversification is relatively rapid, and population sizes are large (Maddison 1997) and so is more likely to mix up closely related organisms than group *Emiliania* with a completely different set of organisms. Furthermore, lineage sorting only acts on genes present in the lineage in question. Genes from red algae, green algae, and heterokonts have to be present to be differentially sorted and lineage sorting does not explain how the haptophytes got these genes.

Horizontal gene transfer from other organisms including endosymbionts seems like the most plausible explanation for the incongruent phylogenetic history of the genes of *Emiliania huxleyi*. Horizontal gene transfer among eukaryotes has historically been viewed as controversial and is widely believed to be less likely in multicellular eukaryotes, particularly animals, in part because the transferred genes need to get into the germ line to be passed on to the next

generation (Keeling and Palmer 2008). Organisms that rely primarily on sexual reproduction are also going to be limited because transferred genes may cause problems during meiotic homologous recombination if only one chromosome has the gene. It has, however, been demonstrated that some animals do have genes from other organisms. One example is the sea slug *Elysia chlorotica*, which has the *psbO* gene from its food *Vaucheria litorea* (Rumpho *et al*. 2008). Another example is the pea aphid, *Acyrthosiphon pisum*, which has acquired genes from its bacterial symbiont, *Buchnera aphidicola* (Nikoh and Nakabachi 2009). Horizontal gene transfer is much more common and less controversial in unicellular eukaryotes (Keeling and Palmer 2008). There is some empirical evidence to support this view. Whitaker *et al*. (2009) found phylogenetic evidence for horizontal gene transfer in several groups of eukaryotes using only trees in which relationships representing putative horizontal gene transfers were supported by bootstrap values of 70% or above. Whitaker *et al*. (2009) found approximately 100 genes that had been transferred to the diatoms, a similar number to the number found in this study for *E. huxleyi*.

Previous studies have implied that some genes are more likely to be transferred than others. Informational genes seem to be transferred less often (Daubin *et al*. 2002; Rivera *et al*. 1998) because of their more complex interactions with other proteins (Jain *et al*. 1999). Recently, it has been proposed that photosynthesis related genes have sufficiently complex interactions to retard the transfer of these genes in a piecemeal manner (Shi and Falkowski 2008). However, these limitations are not absolute and genes with complex interactions

can be transferred horizontally (Wellner *et al*. 2007). Informational and photosynthesis related genes are more likely to be successfully integrated into the new cell if they are transferred in groups of interacting genes. The GeneMerge (Castillo-Davis and Hartl 2003) results of this study indicated that proteins linking *E. huxleyi* to the heterokonts were highly enriched in a few categories of genes, many of them plastid related, while the proteins linking *Emiliania* to the lineages containing primary plastids were enriched in a limited number of categories. This is consistent with the idea that genes, if transferred, are going to be transferred with their functional partners. This study also showed that genes are not evenly distributed on the assembled scaffolds implying that genes grouped by phylogenetic signal may also be grouped physically in the genome providing further evidence that genes were transferred as blocks. For these reasons, HGT seems like the most likely explanation for many of the discordant trees observed in this study, and *Emiliania* appears to be a deeply chimeric organism with a large genetic input from the heterokont, rhodophyte, and maybe Viridiplantae lineages.

Organisms with red-derived secondary plastids

The strongest signal to come from this analysis links *Emiliania huxleyi* to the heterokonts. This is not surprising because haptophytes share many morphological features with heterokonts, despite the lack of mastigonemes that led to their classification as a distinct lineage. Their flagellar roots are similar. Their plastid, which is widely viewed as a secondary plastid of red algal origin, has its stroma surrounded by four membranes and contains chlorophylls a and c.

Mitochondrial cristae are tubular in both groups (Hibberd 1976; Graham and Wilcox 2000; Andersen 2004;). Both groups store carbohydrates as chrysolaminarin (Graham and Wilcox 2000). Because of uncertainty in eukaryotic phylogeny, it is unclear at present which of these are ancestral characters and which might be synapomorphies, but in the bigger picture of eukaryotic phylogeny heterokonts and haptophytes are clearly relatively similar organisms.

One of the more striking features of the heterokonts as a whole is their morphological variability. They include unpretentious flagellates of the genus *Ochromonas*, silica synthesizing diatoms, spectacular giant kelp, and the genus *Phytophthora*, which includes the causative organisms of sudden oak death and potato blight. Recently, *Hyalolithus neolepis*, an organism with morphological and molecular ties to the haptophytes, was found to make its scales from silica, which demonstrates that at least one haptophyte can metabolize this mineral (Yoshida *et al*. 2006). Some previous studies have found a relationship between haptophytes and heterokonts based on plastid genes with moderate to strong support (Yoon *et al*. 2002; Yoon *et al*. 2004; Bachvaroff *et al*. 2005).

In contrast, several studies have found evidence that haptophytes are most closely related to cryptophytes. Le Corguillé *et al*. (2009) used plastid genes to find a relationship between *Emiliania huxleyi* and cryptophytes with moderate to strong support. Riisberg *et al*. (2009) used ribosomal, cytosolic protein coding genes, and *rbcL* to show a strongly supported sister relationship between haptophytes and cryptophytes. Other groups have found haptophytes

and cryptophytes to be monophyletic using multiple nuclear genes (Burki *et al.* 2007; Hackett *et al.* 2007). My pipeline found only four sequences that supported a sister relationship between the haptophytes and cryptophytes. This could be due to the smaller amount of data provided by EST collections compared to complete genomes. The four translated genes, when subjected to phylogenetic analysis, did strongly support the grouping of *E. huxleyi* with *Guillardia theta*. It is possible that haptophytes and cryptomonads really are sister taxa, but more data are needed before that statement can be made with confidence. These studies (Burki *et al*. 2007; Hackett *et al*. 2007) also strongly linked the alveolates with the heterokonts to the exclusion of the haptophytes and cryptophytes. My analysis found 55 proteins that indicate that alveolates and heterokonts form a monophyletic group to the exclusion of the haptophytes and cryptophytes. This is still far fewer than the 305 proteins that put *E. huxleyi* in the heterokonts. This suggests that the *E. huxleyi* heterokont ancestral linkage should be reconsidered along with the current configuration of the chromalveolates in general. It is interesting that the four proteins found in this study that put *E. huxleyi* as sister to *Guillardia theta* showed this pair as being close to *Dictyostelium discoidium* and the opisthokonts similar to the placement found by Bhattacharya *et al*. (1993). This study supports the monophyly of chromalveolates (alveolates, heterokonts, haptophytes, and cryptophytes) in general but the bulk of the data in this study put *Emiliania* within the heterokonts not with the cryptophytes as other studies have suggested (Burki *et al* 2007; Hackett *et al* 2007).

Emiliania and the primary plastid containing lineages

In addition to the protein trees linking *Emiliania* to the heterokonts, this study revealed a large number of trees linking *E. huxleyi* to each of the two primary plastid containing lineages. Somewhat surprisingly, there were a larger number of protein trees linking *E. huxleyi* with green lineage organisms than with red lineage organisms. Moustafa *et al*. (2009) also found a large number of green lineage derived genes in diatoms and *E. huxleyi*. They dismissed the hypothesis that this observation was an analytical artifact, and offered two possible explanations: 1) Heterokonts had a green algal endosymbiont before they had a red algal endosymbiont. This explanation requires two plastid gains and one plastid loss. 2) The ancestral host heterokont cell was actually an *Ostreococcus*-like green alga that lost its plastid then gained a red algal endosymbiont. This explanation requires a plastid loss followed by a plastid gain. There is ample evidence for multiple plastid gains and losses in dinoflagellates (Nosenko *et al*. 2006; Patron *et al*. 2006) so the hypotheses of Moustafa *et al*. (2009) are not only possible but very interesting and worth continued consideration. While the results of the present study found more trees linking *Emiliania* with the green lineage organisms as a whole, there were more complete genomes available for green lineage organisms. *E. huxleyi* more frequently grouped with *Cyanidioschyzon* than any particular group of green lineage organisms. This implies that as more rhodophyte genomes become available, the red signal will become stronger than the green signal and that what seems like a green signal now may actually be a primary plastid containing

lineage signal. In addition, this study found almost as many proteins linking *E. huxleyi* with the heterokonts and the rhodophytes together as with the heterokonts and the Viridiplantae together despite the larger amount of genomic data available from the Viridiplantae (fig 4). Furthermore, when these alignments were concatenated, all of the resulting trees put *E. huxleyi* with the heterokonts. The full dataset *E. huxleyi*, heterokont, rhodophyte tree significantly supported a relationship between the *E. huxleyi*-heterokont bipartition and the rhodophyte *Cyanidioschyzon merolae*. The trees resulting from concatenation of protein alignments that individually supported an *E. huxleyi*-heterokont-Viridiplantae bipartion significantly supported a relationship between *E. huxleyi* and the heterokonts but did not support a sister relationship between the *E. huxleyi*-heterokont bipartition and the green lineage. The full dataset *E. huxleyi*, heterokont, Viridiplantae tree actually put rhodophytes with the heterkont-*E. huxleyi* bipartition. This would support the more conventional idea that the originally aplastidic chromists gained one secondary plastid of red algal origin. This should also be considered the more parsimonious explanation because it only requires one plastid gain and no losses.

Who is the ancestor?

The two lineages with large numbers of proteins linking them to *Emiliania huxleyi* are the heterokonts and the Viridiplantae. Of these, more genes linked the heterokonts to the *E. huxleyi*. These genes appear to be enriched in plastid related functions indicating that the plastids of haptophytes and heterokonts are related. This is not controversial as the plastids share many characteristics

including the presence of chlorophylls a and c, thylakoids in stacks of three, and four bounding membranes around the stroma (Andersen 2004). This just means the plastids are related, but not necessarily the host organisms. It is possible that the heterokonts and haptophytes got their plastids from a common ancestor, but it is still possible that they acquired their plastids in separate events from a similar source. In these scenarios, the apparently green derived genes could be part of a larger set of genes from the primary plastid-containing lineage that will turn out to be from rhodophytes as more red algal genomes become available. Non-plastid related genes could have been transferred from the plastid donor and the plastid related genes diverged more rapidly after being incorporated into the new host genome. This seems sensible since the plastids of haptophytes and heterokonts are quite different from the plastids of red algae in that red algae have phycobiliproteins but lack chlorophyll c and only have two membranes around the stroma (Graham and Wilcox 2000). It has also been demonstrated that plastid genes can evolve quite rapidly in dinoflagellates although the same study indicated that plastid genes change more slowly in *Emiliania* (Bachvaroff *et al*. 2006).

Emiliania seems to have another set of genes from the Viridiplantae. These genes, though less numerous, seem to have more varied function implying that maybe the host organism descended from a green algae like organism. Other investigators have suggested this in various forms. Cavalier-Smith has suggested that haptophytes and heterokonts descended from a Viridiplantae-like organism (Cavalier-Smith 1981, Cavalier-Smith 1999). More

recently Burki *et al*. (2008) published a concatenated tree that placed most photosynthetic lineages as sharing a common ancestor. The haptophytes were closer to the primary plastid lineages than the heterokonts, consistent with my finding that *Emiliania* has a set of green derived genes that it does not share with the heterokonts. Moustafa *et al*. (2009) found a large number of green genes in heterokonts and *Emiliania.* This is also consistent with the heterokonts and haptophytes having descended from a green alga. However, when I concatenated proteins that individually grouped *Emiliania* with the heterokonts and the green lineage, the resulting tree did not show a significant relationship between the green lineage and the heterokont-*E. huxleyi* bipartition.

Conclusions

It was not a surprise that my pipeline revealed that *Emiliania huxleyi* has genetic elements from multiple sources. Coccoliths have been confidently identified from Triassic sediments (Fawlkowski *et al*. 2004) indicating the coccolith bearing haptophytes have been around for at least 200 million years and the haptophytes that do not make coccoliths may be even more ancient. Haptophytes are plastidic and can be phagotrophic (Kawachi *et al*. 1991) giving them both the time and the mechanism to acquire horizontally transferred genes. The pipeline indicated a large number of proteins linking *Emiliania* phylogenetically to the heterokonts and other proteins to the green algae and the red algae. The distribution of the proteins associating *E. huxleyi* with the heterokonts and the green lineage differed from the distribution predicted based on the length of each scaffold indicating that genes from different sources are

clustered physically in the *E. huxleyi* genome. This physical clustering of the different genomic inputs provides further evidence for the chimeric nature of the *Emiliania* genome because large groups of genes could have been transferred together in a small number of transfer events. Genes from different sources also tended to have related functions. Proteins that grouped *E. huxleyi* with the heterokonts had primarily plastid related functions. Proteins indicating a relationship between *Emiliania* and the green lineage were more varied functionally although they still fell into a small number of GO categories. Thus genes were imported with their functional partners making it easier for the proteins to function in their new context. The larger number of genes linking *E. huxleyi* to the heterokonts implies that *E. huxleyi* may be more closely related to heterokonts. The more varied function implies that *E. huxleyi* may be closer to the Viridiplantae. The multiprotein tree based on individual proteins that indicated a relationship between *Emiliania* and the green lineage did support a relationship between the green lineage and *E. huxleyi*. However, when proteins that individually grouped *E. huxleyi* with the heterokonts and the Viridiplantae analyzed together, *Emiliania* grouped with the heterokonts and the relationship with the green lineage did not receive significant support. Furthermore, when proteins that linked *E. huxleyi* to the heterokonts and the rhodophytes were concatenated, *Emiliania* again fell out with the heterokonts, but the relationship of this group with the rhodophyte *Cyanidioschyzon merolae* did receive significant support. Most of the evidence from this study indicates that *E. huxleyi* is closer

to the heterokonts. More data from cryptophytes and rhodophytes are needed to understand the evolutionary history of the chromalveolates.

Chapter IV: Conclusions

Symbiosis is widely believed to date back to the origins of eukaryotic life (Cavalier-Smith 2006; Margulis 1996; Martin and Russell 2003). Mitochondria (Grey *et al*. 1999) and plastids (Reith 1995) are both regarded as being of bacterial origin. Symbiosis can be divided into three major types: mutualism, commensalism, and parasitism (Saffo 1993). Mutualism involves two organisms working together and helping each other. Commensalism involves one organism not being harmed by the organism it is helping. Parasitism involves a relationship in which one organism is taking advantage of and harming the other. I have addressed parasitism and mutualism in my dissertation research. *Amoebophrya* enters the cytoplasm of a host, grows, then leaves the host dead (Cachon and Cachon 1987) clearly harming the host. *Emiliania huxleyi* is a haptophyte phytoplankton with a secondary plastid of red algal origin (Delwiche 1999). Secondary plastids are plastids that resulted from the engulfment of a phototrophic eukaryote by another eukaryote (Palmer and Delwiche 1996). Modern examples of phototrophs living within other organisms without having been fully incorporated into the host cell as an organelle indicate that these relationships are mutualistic. The host benefits from photosynthate that leaks from the algal symbiont and the algae benefit from host metabolic byproducts the symbionts use as nutrients (Yellowlees *et al*. 2008).

Amoebophrya belongs to a group of dinoflagellate-like organisms referred to as group II marine alveolates (López-Garcia *et al*. 2001). Published results indicate that group II marine alveolates are among the most diverse eukaryotic

organisms in enclosed marine habitats (Cheung *et al*. 2008) and the open ocean even at depths of 3000 m (Groisillier *et al*. 2006; López-Garcia *et al*. 2001; Not *et al*. 2007). Group II alveolate sequences have also been found in hypersaline anoxic waters (Alexander *et al*. 2009) and at hydrothermal sediments (López-Garcia *et al*. 2003). Despite their apparent abundance and diversity, to my knowledge the only group II alveolate genera that have been described morphologically are *Syndinium* (Manier 1971; Soyer 1974), *Hematodinium* (Appleton and Vickerman 1998), and *Amoebophrya* (Cachon and Cachon 1987). I have added to the body of knowledge about this underdescribed group in several ways. 1) I described the free-swimming stage of the *Amoebophrya* life cycle at an ultrastructural level. 2) I did experiments with cytochalasin D that indicate *Amoebophrya* uses a microfilament-based mechanism to enter its host. 3) I provided ultrastructural evidence that the mastigocoel is formed as a result of vesicle fusion rather than a sheet of cytoplasm folding around a central column of cytoplasm. Despite our knowledge of their diversity and my additions to the knowledge about *Amoebophrya* structure and function there is still a lot to learn about group II marine alveolates. All of the morphologically described group II alveolates are parasites implying that many if not all group II alveolates may be parasites. Their parasitic nature may be the reason that so few organisms from this group have been isolated. Parasitism is thought to be a good strategy in high productivity environments, but low productivity environments favor mutualism (Thrall *et al*. 2006). Yet group II marine alveolates are found in low productivity environments below the photic zone or even in anaerobic

environments where finding a new host is likely to be difficult. Are these organisms still highly virulent parasites or do they form more stable symbioses in low productivity environments?

Emiliania huxleyi is a haptophyte alga, which possesses a secondary plastid containing chlorophylls a and c. Haptophytes have traditionally been linked to heterokonts (Andersen 2004). They were placed within the chromalveolates because they have chlorophyll a and c containing secondary plastids, as do some members of each of the other groups including the alveolates, heterokonts, and cryptomonads (=cryptophytes) (Cavalier-Smith 1999). Recent studies have placed haptophytes as sister to the cryptomonads (Burki *et al*. 2007; Patron *et al*. 2007). My phylogenomic pipeline detected genetic signals from three sources: heterokont, green algae, and red algae within the genome of *E. huxleyi.* This supports the idea that *E. huxleyi* is a chimeric organism, but the apparent presence of green algal genes seems at odds with the conventional wisdom that haptophyte plastids are of red algal origin. Further evidence that the *Emiliania* genome is chimeric was provided by Gene Merge results which demonstrated that the genes linking the *E. huxleyi* to the heterokonts and the greens each were enriched in their own small set of functional categories implying that genes were transferred with functional partners. Another set of evidence supporting the chimeric nature of *Emiliania* is my finding that genes are distributed on scaffolds differently than would be expected based on scaffold length, implying that genes were transferred as large multigene units. Moustafa *et al*. (2009) found evidence that diatoms, a type of

heterokont, also have genes of green algal origin. They suggested that either the diatoms descended from a green alga or that they had harbored a green algal symbiont prior to acquiring the red algal symbiont. Either of these scenarios is possible although it is also worth mentioning that more green algal genomes than red algal genomes have been sequenced. It is possible that as more red algal genomes are sequenced the genes apparently from green algae will turn out to have been from a red algal source. At this point the genetic input from red algal and green algal sources should really be considered together as a Viridiplantae contribution. My pipeline only identified four proteins that linked *Emiliania* to the cryptomonad *Guillardia*. When these proteins were concatenated and another phylogenetic analysis was done, they strongly supported a sister relationship between *Emiliania* and *Guillardia*. However, the number of proteins supporting this relationship is much lower than the number of proteins supporting a relationship between *E. huxleyi* and the heterokonts or the primary plastid containing lineages. This may be due to a lack of sequence data. At present there are no completely sequenced cryptomonad genomes. To resolve these issues more red algal and cryptomonad complete genome sequences are needed.

I have explored symbiosis both from the standpoint of a parasite living in a host cell and from the standpoint an organism that harbors an endosymbiotic organelle. Being an intracellular parasite and having a plastid may seem very different at first and it may be surprising that two chromalveolate groups, the heterokonts and the alveolates, include both phototrophs and parasites, but the

difference between being an endosymbiont and having an endosymbiont may not be as great as first appears. If you imagine two equal sized and equally capable amoeboid cells, each of which will try to consume anything in its path, crossing paths, assuming everything else is equal, it is actually a matter of chance which cell ends up inside. Some of the requirements for having a plastid and for being an intracellular parasite are similar. Both need to be able to transport gene products into a foreign cell. It has now been demonstrated that special leader sequences are necessary to target nucleus-encoded proteins to plastids in apicomplexans (Harb *et al*. 2004; Waller *et al*. 1998), dinoflagellates (Nassoury *et al*. 2003; Patron *et al*. 2005), diatoms (Gruber *et al*. 2007; Lang *et al*. 1998), and a cryptomonad (Gould *et al*. 2006). The leader sequences of chromalveolates typically have two parts. The outer membrane of the plastid is part of the endoplasmic reticulum and there is a signal peptide that targets plastid directed proteins to and through that membrane. There is also a transit peptide that takes the protein into the plastid itself (Lang *et al*. 1998; Nassoury *et al*. 2003; Waller *et al*. 1998). Chromalveolate parasites such as *Plasmodium* and *Phytophthora* need to target proteins into the cytoplasm of their host cells. To do this they use a signal sequence to get the protein into the secretory pathway and to the parasitophorous vacuole compartment. They then use an additional vacuolar translocation sequence to export proteins to the cytoplasm of the host cell (Haldar *et al*. 2006). The presequence required to get a protein into a host cell cytoplasm is similar to the presequence needed to get a protein into a secondary

plastid. The possession of a plastid could have predadapted these organisms to become parasites and made them more attractive symbionts.

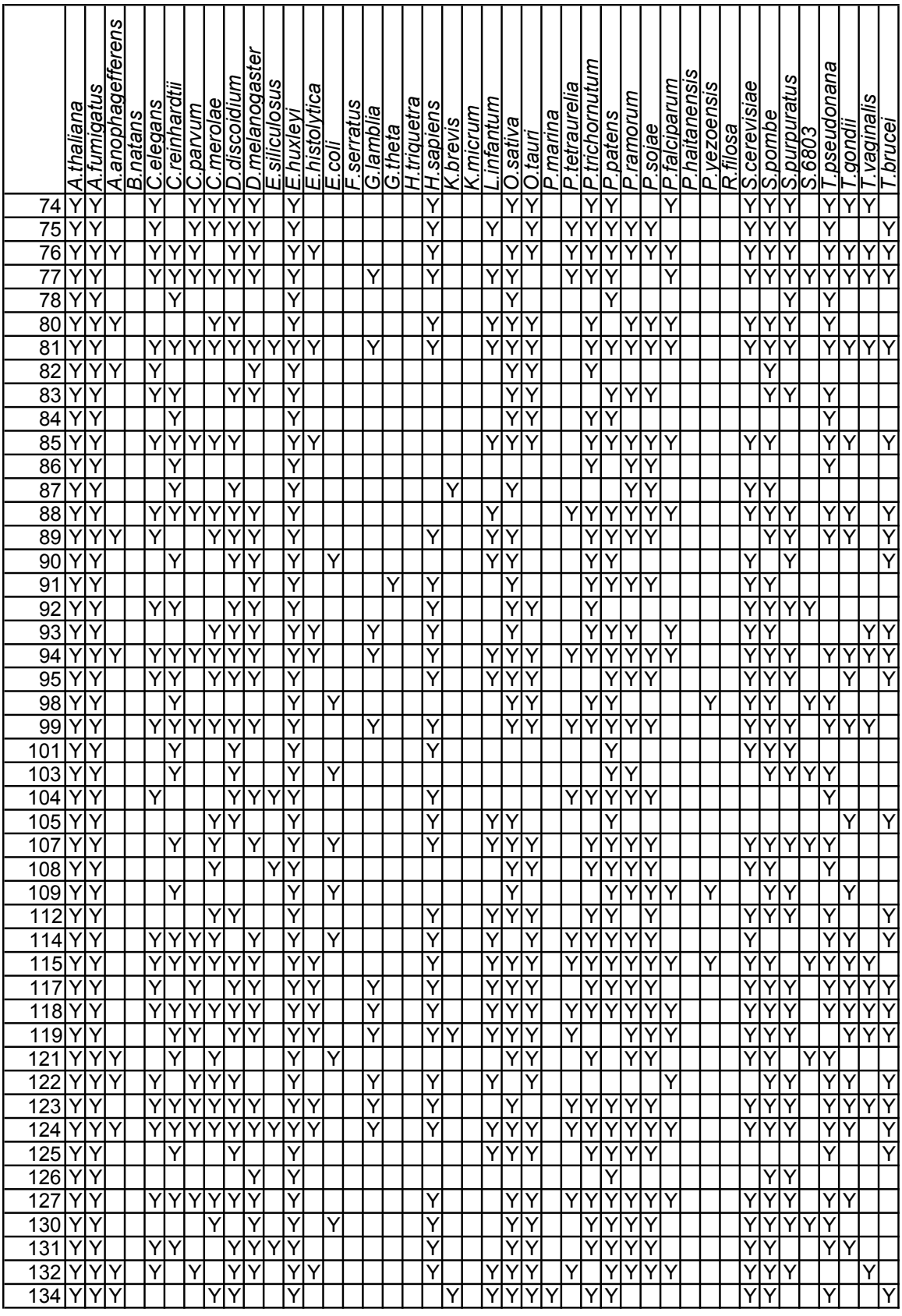
The presence of *Chromera,* a phototrophic endosymbiont, at the base of the apicomplexan phylogenetic tree (Moore *et al*. 2008) also argues for parasites, at least in some organisms, starting as endosymbiotic phototrophs. As I mentioned above, low productivity environments tend to favor mutualism while high productivity environments tend to favor parasitism (Thrall *et al*. 2006). Reef building corals that harbor *Symbiodinium* tend to live in nutrient poor environments (Muscatine and Porter 1977). *Chromera* was isolated from *Plesiastrea versipora* a scleractinian coral (Moore *et al*. 2008) that is widespread in coral reefs (Rodriguez-Lanetty and Hoegh-Guldberg 2002). Thus *Chromera* most likely frequents oligotrophic, nutrient poor, waters. As apicomplexans and possibly group II marine alveolates invaded more productive waters and terrestrial environments, that favor thicker and frequently less translucent host ectoderms, they transitioned from mutualism to parasitism. Given that *Amoebophrya* (Chapter 2) and oomycetes including *Phytophthora* (Hardham 2007) seem to use electron dense vesicles to enter host cells despite lacking an obvious apical complex, it would be interesting to know if *Chromera* has some type of a cryptic apical complex consisting of electron dense structures.

Chromalveolates are a diverse set of organisms that provide many interesting opportunities to study symbiosis. I have studied the parasitic group II alveolate *Amoebophrya* and have shown that the free-swimming dinospores have large numbers of electron dense vesicles, which seem to line up in a tube

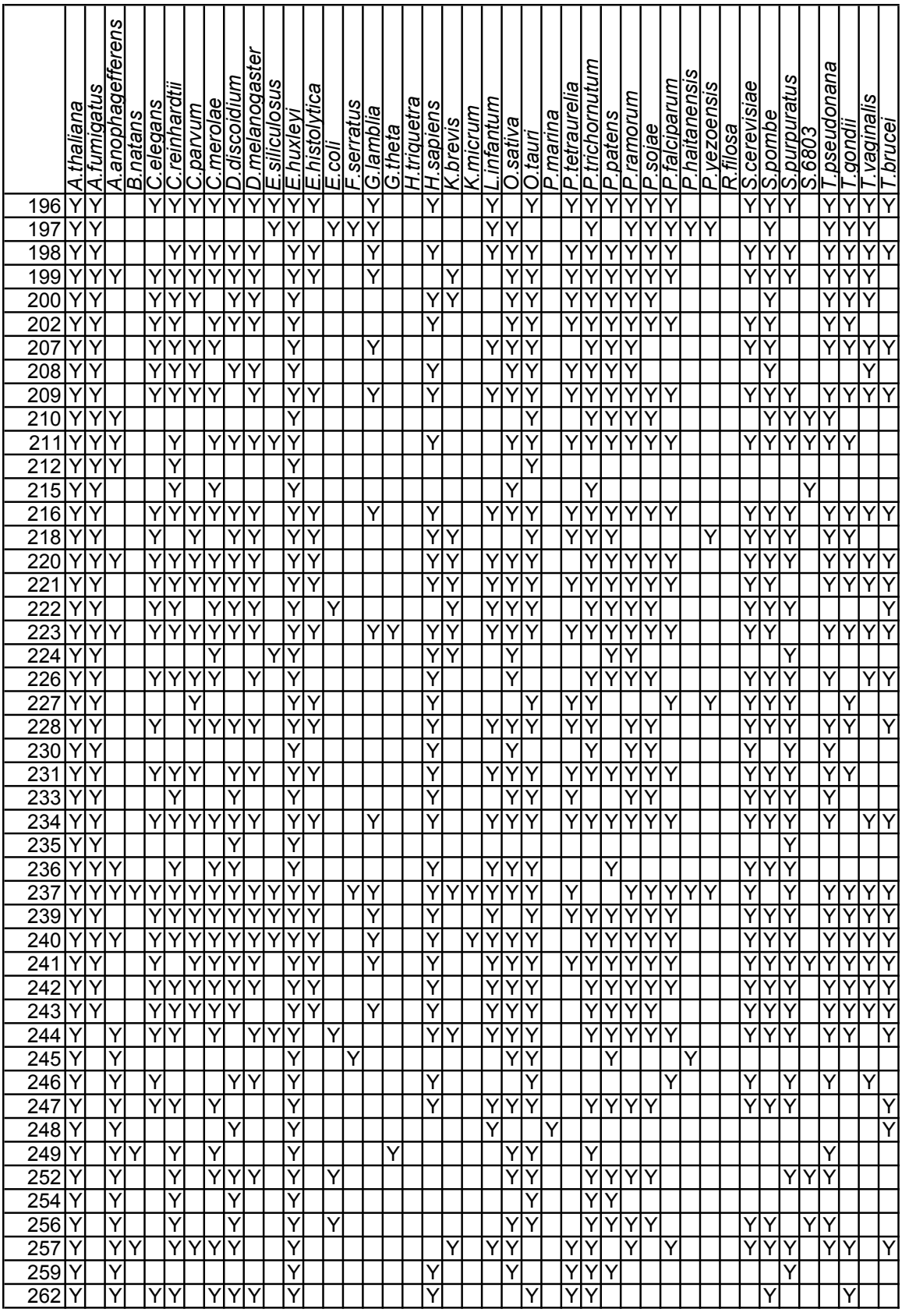
as the parasite is attached to the host surface. Cytochalasin experiments I did imply that *Amoebophrya* uses an actin based mechanism to enter host cells. These electron dense organelles and the use of microfilaments in host entry underscores that *Amoebophrya* is apicomplexan-like as well as dinoflagellate-like. I also provided evidence that the internal compartment of mature *Amoebophrya* trophonts called the mastigocoel is formed as a result of vesicle fusion. My study of the *Emiliania huxleyi* genome revealed that it contains large numbers of genes that link it phylogenetically to heterokonts and to primary plastid containing organisms emphasizing its chimeric nature. Clearly, there is need for additional ultrastructural and genomic work on these and other abundant, diverse, and relatively understudied protists.

Appendix: Taxonomic contents of COGs

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichornutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
5	Y	Y			Y		Y		Y	Y		Y	Y						Y			Y	Y		Y		Y		Y								Y		Y	Y		
6	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y						Y			Y		Y	Y	Y	Y		Y	Y					Y	Y	Y		Y	Y	Y	Y
9	Y	Y	Y						Y			Y	Y										Y				Y		Y	Y									Y			
11	Y	Y			Y	Y			Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y						Y	Y		Y		Y	Y
12	Y	Y				Y			Y	Y		Y											Y	Y			Y	Y	Y	Y	Y					Y	Y	Y	Y	Y		
13	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y		Y		Y	Y			Y	Y	Y	Y
14	Y	Y	Y		Y	Y	Y		Y	Y		Y	Y						Y				Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y			Y		
15	Y	Y			Y	Y	Y		Y	Y		Y							Y				Y	Y		Y		Y	Y	Y						Y	Y			Y		
16	Y	Y	Y		Y		Y		Y	Y		Y				Y			Y	Y		Y		Y		Y	Y		Y	Y	Y				Y	Y	Y		Y		Y	Y
17	Y	Y			Y							Y	Y	Y		Y			Y								Y								Y							
18	Y	Y						Y	Y		Y	Y											Y	Y			Y	Y	Y	Y					Y	Y		Y	Y			
19	Y	Y			Y	Y	Y	Y	Y			Y	Y									Y	Y	Y		Y	Y		Y	Y	Y				Y	Y	Y		Y	Y		Y
20	Y	Y			Y				Y	Y		Y							Y				Y	Y		Y	Y	Y							Y	Y	Y		Y	Y	Y	
21	Y	Y			Y	Y		Y		Y		Y							Y								Y	Y							Y	Y			Y			
22	Y	Y			Y	Y	Y	Y	Y	Y		Y	Y							Y		Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
23	Y	Y			Y	Y	Y		Y	Y		Y	Y						Y				Y			Y	Y	Y	Y	Y					Y	Y	Y		Y	Y		
25	Y	Y						Y				Y		Y									Y				Y	Y	Y						Y	Y		Y	Y			
27	Y	Y	Y		Y	Y	Y	Y	Y	Y		Y							Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y			Y	Y		Y
28	Y	Y	Y		Y	Y	Y		Y	Y		Y							Y				Y					Y	Y							Y	Y			Y	Y		
30	Y	Y			Y				Y	Y		Y							Y				Y					Y		Y							Y			Y		
31	Y	Y			Y	Y		Y	Y	Y	Y	Y							Y			Y	Y	Y		Y	Y	Y	Y	Y	Y		Y		Y	Y	Y		Y		Y	Y
32	Y	Y			Y	Y		Y	Y	Y	Y	Y		Y					Y				Y	Y			Y	Y	Y	Y	Y				Y	Y	Y	Y	Y			
33	Y	Y	Y		Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
34	Y	Y			Y		Y	Y	Y	Y		Y	Y			Y			Y			Y		Y		Y	Y	Y	Y	Y					Y	Y	Y		Y		Y	Y
36	Y	Y			Y	Y	Y		Y	Y		Y							Y				Y	Y		Y		Y		Y					Y	Y			Y			
40	Y	Y				Y			Y	Y		Y	Y									Y	Y	Y			Y	Y		Y					Y	Y			Y			Y
43	Y	Y	Y		Y	Y	Y	Y	Y		Y	Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y			Y	Y	Y	Y
44	Y	Y	Y		Y	Y		Y	Y	Y		Y		Y					Y				Y	Y			Y	Y	Y	Y	Y				Y	Y	Y	Y	Y	Y		Y
45	Y	Y			Y	Y		Y	Y	Y		Y							Y				Y			Y	Y	Y	Y	Y				Y			Y		Y			
47	Y	Y	Y		Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y		Y	Y	Y				Y	Y	Y		Y	Y		Y
49	Y	Y			Y		Y	Y	Y	Y		Y							Y			Y	Y			Y		Y	Y	Y	Y				Y	Y	Y		Y	Y		Y
51	Y	Y	Y					Y	Y			Y											Y	Y		Y		Y		Y					Y	Y						
52	Y	Y			Y	Y			Y		Y	Y		Y					Y			Y	Y			Y	Y	Y	Y	Y							Y		Y			
55	Y	Y	Y		Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y			Y	Y	Y	Y	Y		Y		Y	Y	Y		Y	Y	Y	Y
56	Y	Y			Y		Y	Y	Y	Y	Y	Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y			Y	Y	Y	Y
57	Y	Y								Y		Y							Y			Y	Y					Y								Y				Y		Y
59	Y	Y			Y	Y			Y	Y		Y	Y						Y				Y	Y				Y	Y	Y					Y	Y	Y				Y	Y
60	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y							Y			Y	Y				Y	Y	Y		Y					Y	Y		Y			Y
61	Y	Y						Y				Y																														
62	Y	Y			Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
63	Y	Y			Y	Y	Y	Y	Y	Y		Y	Y						Y			Y	Y	Y			Y	Y	Y	Y					Y	Y	Y		Y			Y
64	Y	Y	Y			Y		Y				Y		Y									Y	Y			Y	Y	Y						Y	Y		Y	Y			
65	Y	Y			Y		Y			Y		Y							Y				Y	Y			Y		Y	Y	Y				Y	Y	Y			Y		
66	Y	Y			Y	Y		Y	Y	Y	Y	Y				Y			Y				Y				Y		Y	Y	Y					Y	Y		Y	Y		Y
67	Y	Y			Y	Y	Y	Y	Y	Y		Y	Y						Y				Y	Y		Y		Y	Y	Y					Y	Y			Y	Y	Y	
70	Y	Y							Y			Y											Y					Y							Y	Y						
71	Y	Y			Y	Y			Y	Y		Y	Y						Y			Y	Y	Y				Y	Y	Y					Y	Y	Y				Y	Y


A.thaliana
A.fumigatus
A.anophagefferens
B.natans
C.elegans
C.reinhardtii
C.parvum
C.merolae
D.discoidium
D.melanogaster
E.siliculosus
E.huxleyi
E.histolytica
E.coli
F.serratus
G.lamblia
G.theta
H.triquetra
H.sapiens
K.brevis
K.micrum
L.infantum
O.sativa
O.tauri
P.marina
P.tetraurelia
P.trichornutum
P.patens
P.ramorum
P.sojae
P.falciparum
P.haitanensis
P.yezoensis
R.filosa
S.cerevisiae
S.pombe
S.purpuratus
S.6803
T.pseudonana
T.gondii
T.vaginalis
T.brucei
74
75
76
77
78
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
98
99
101
103
104
105
107
108
109
112
114
115
117
118
119
121
122
123
124
125
126
127
130
131
132
134

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichornutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
136	Y	Y			Y	Y	Y	Y	Y			Y	Y						Y			Y	Y	Y		Y		Y	Y	Y				Y	Y	Y	Y		Y	Y	Y	
137	Y	Y				Y						Y											Y				Y	Y	Y					Y					Y			
138	Y	Y			Y						Y	Y				Y										Y				Y					Y			Y			
139	Y	Y			Y	Y	Y	Y		Y		Y							Y			Y	Y	Y		Y	Y	Y	Y	Y				Y				Y			Y
140	Y	Y	Y		Y	Y	Y	Y	Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y	Y			Y	Y	Y		Y		Y	Y
142	Y	Y	Y		Y	Y	Y	Y	Y		Y	Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y			Y	Y	Y		Y	Y	Y	Y
144	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y				Y		Y			Y		Y		Y	Y		Y	Y	Y			Y	Y	Y		Y	Y	Y	Y
145	Y	Y			Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y		Y	Y	Y			Y	Y			Y	Y	Y	Y
146	Y	Y			Y			Y	Y	Y		Y							Y	Y		Y	Y	Y		Y	Y	Y	Y	Y	Y			Y	Y	Y		Y			Y
147	Y	Y			Y	Y	Y	Y	Y	Y		Y	Y	Y		Y			Y			Y	Y	Y			Y	Y	Y	Y	Y		Y		Y	Y	Y	Y	Y	Y	Y	Y
148	Y	Y			Y	Y	Y		Y	Y		Y	Y						Y				Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	
150	Y	Y			Y	Y		Y	Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y					Y	Y			Y	Y	Y	
151	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y					Y			Y		Y		Y			Y		Y				Y	Y	Y			Y		Y
152	Y	Y							Y			Y																							Y							
154	Y	Y			Y	Y	Y	Y	Y	Y		Y	Y			Y			Y	Y		Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
155	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y			Y			Y				Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
156	Y	Y			Y	Y	Y		Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
157	Y	Y				Y		Y	Y	Y		Y			Y				Y					Y				Y	Y	Y		Y				Y						
158	Y	Y				Y	Y			Y		Y							Y			Y	Y	Y			Y	Y							Y				Y	Y		Y
159	Y	Y					Y					Y		Y															Y	Y											
160	Y	Y			Y			Y		Y		Y							Y				Y	Y			Y		Y	Y					Y	Y	Y	Y	Y			
161	Y	Y	Y		Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
164	Y	Y			Y		Y	Y	Y	Y		Y	Y						Y			Y		Y		Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
165	Y	Y					Y	Y	Y			Y				Y						Y				Y	Y		Y	Y				Y	Y	Y				Y		
166	Y	Y			Y	Y			Y	Y		Y							Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y		Y
167	Y	Y										Y																						Y								
168	Y	Y							Y	Y		Y							Y					Y		Y			Y	Y					Y	Y	Y					
170	Y	Y	Y		Y	Y			Y	Y	Y	Y			Y				Y	Y		Y	Y	Y					Y	Y			Y		Y	Y	Y		Y			Y
171	Y	Y							Y			Y		Y					Y				Y	Y			Y	Y						Y	Y							
172	Y	Y	Y		Y			Y	Y	Y		Y							Y				Y						Y	Y				Y						Y		
173	Y	Y			Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y			Y	Y	Y	Y				Y	Y			Y			Y
174	Y	Y			Y					Y		Y							Y				Y				Y	Y	Y	Y				Y	Y	Y		Y			
175	Y	Y			Y		Y	Y	Y	Y		Y	Y						Y			Y		Y			Y	Y	Y	Y	Y			Y	Y	Y		Y	Y	Y	Y
176	Y	Y										Y										Y		Y			Y												Y		Y	
178	Y	Y			Y	Y		Y	Y	Y		Y		Y					Y	Y	Y	Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y	Y	Y	Y		Y
179	Y	Y			Y	Y	Y	Y	Y	Y		Y							Y			Y	Y	Y		Y		Y	Y	Y	Y				Y	Y			Y	Y		Y
180	Y	Y			Y			Y	Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y					Y		Y					Y
182	Y	Y	Y			Y	Y	Y	Y	Y	Y	Y	Y						Y			Y	Y	Y			Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
183	Y	Y			Y	Y	Y	Y	Y	Y		Y	Y						Y				Y	Y					Y						Y	Y	Y		Y		Y	
184	Y	Y			Y	Y			Y	Y		Y		Y								Y		Y			Y	Y	Y	Y	Y				Y	Y	Y	Y	Y	Y		Y
185	Y	Y							Y			Y										Y	Y					Y							Y	Y			Y			Y
187	Y	Y										Y																Y							Y					Y		
190	Y	Y			Y	Y	Y	Y	Y	Y		Y															Y		Y	Y	Y				Y	Y	Y		Y	Y		
192	Y	Y		Y	Y			Y	Y			Y		Y					Y			Y	Y				Y		Y	Y					Y	Y	Y	Y	Y	Y		Y
193	Y	Y	Y		Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
194	Y	Y				Y		Y				Y		Y									Y	Y		Y	Y	Y	Y						Y	Y		Y	Y			
195	Y	Y			Y			Y	Y	Y		Y	Y	Y									Y	Y		Y	Y	Y	Y	Y						Y	Y		Y			


A.thaliana
A.fumigatus
A.anophagefferens
B.natans
C.elegans
C.reinhardtii
C.parvum
C.merolae
D.discoidium
D.melanogaster
E.siliculosus
E.huxleyi
E.histolytica
E.coli
F.serratus
G.lamblia
G.theta
H.triquetra
H.sapiens
K.brevis
K.micrum
L.infantum
O.sativa
O.tauri
P.marina
P.tetraurelia
P.trichornutum
P.patens
P.ramorum
P.sojae
P.falciparum
P.haitanensis
P.yezoensis
R.filosa
S.cerevisiae
S.pombe
S.purpuratus
S.6803
T.pseudonana
T.gondii
T.vaginalis
T.brucei
196
197
198
199
200
202
207
208
209
210
211
212
215
216
218
220
221
222
223
224
226
227
228
230
231
233
234
235
236
237
239
240
241
242
243
244
245
246
247
248
249
252
254
256
257
259
262

	A.thaliana	A.fumigatus	A.anophagefferens	B.natans	C.elegans	C.reinhardtii	C.parvum	C.merolae	D.discoidium	D.melanogaster	E.siliculosus	E.huxleyi	E.histolytica	E.coli	F.serratus	G.lamblia	G.theta	H.triquetra	H.sapiens	K.brevis	K.micrum	L.infantum	O.sativa	O.tauri	P.marina	P.tetraurelia	P.trichornutum	P.patens	P.ramorum	P.sojae	P.falciparum	P.haitanensis	P.yezoensis	R.filosa	S.cerevisiae	S.pombe	S.purpuratus	S.6803	T.pseudonana	T.gondii	T.vaginalis	T.brucei
266	Y		Y						Y			Y							Y				Y			Y	Y	Y	Y	Y				Y	Y			Y	Y		
267	Y		Y			Y		Y			Y	Y							Y								Y		Y	Y	Y				Y	Y	Y		Y		Y	
273	Y		Y									Y											Y				Y	Y	Y	Y					Y	Y			Y			
274	Y		Y		Y	Y			Y	Y	Y	Y	Y						Y				Y	Y	Y	Y		Y	Y	Y					Y	Y	Y				Y	Y
275	Y		Y		Y		Y	Y				Y	Y				Y		Y			Y		Y			Y						Y				Y			Y	Y
277	Y		Y			Y						Y																Y	Y	Y											
278	Y		Y		Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
281	Y		Y		Y	Y			Y	Y		Y		Y					Y				Y	Y		Y	Y	Y	Y	Y					Y	Y	Y		Y	Y		Y
284	Y		Y		Y	Y			Y		Y	Y							Y			Y	Y			Y	Y								Y	Y	Y		Y			
285	Y		Y		Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y								Y	Y	Y		Y		Y	Y
286	Y		Y			Y	Y	Y	Y		Y	Y									Y	Y	Y	Y		Y	Y	Y	Y	Y		Y							Y			
287	Y		Y		Y	Y	Y	Y	Y	Y	Y	Y	Y			Y			Y			Y		Y		Y	Y	Y	Y	Y					Y	Y		Y	Y	Y	Y	Y
288	Y		Y			Y			Y			Y											Y	Y		Y	Y	Y	Y	Y									Y	Y		Y
290	Y		Y			Y	Y	Y	Y	Y	Y	Y	Y			Y			Y	Y		Y	Y	Y		Y	Y	Y	Y	Y					Y	Y			Y	Y	Y	Y
291	Y		Y		Y	Y		Y	Y	Y		Y					Y		Y			Y	Y				Y	Y	Y	Y							Y		Y	Y		Y
294	Y		Y			Y		Y		Y		Y							Y				Y	Y				Y										Y		Y			
295	Y		Y		Y	Y	Y	Y	Y	Y	Y	Y	Y			Y						Y	Y			Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
300	Y		Y		Y							Y																Y											Y			
308	Y		Y									Y											Y	Y				Y	Y													
310	Y		Y			Y					Y	Y						Y					Y	Y			Y	Y							Y	Y	Y		Y			
316	Y		Y		Y	Y	Y		Y	Y		Y				Y			Y			Y	Y	Y		Y	Y	Y	Y	Y							Y		Y		Y	
317	Y		Y		Y		Y	Y	Y	Y		Y	Y	Y		Y			Y			Y		Y		Y		Y		Y	Y				Y	Y				Y	Y	Y
318	Y		Y		Y	Y		Y	Y	Y		Y							Y	Y		Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y		
319	Y		Y		Y	Y	Y	Y	Y	Y		Y	Y	Y		Y	Y		Y		Y	Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y	Y	Y	Y	Y	Y
321	Y		Y			Y		Y	Y	Y	Y	Y		Y					Y			Y	Y	Y		Y		Y	Y	Y	Y				Y	Y		Y	Y	Y		
323	Y		Y								Y	Y					Y										Y															
324	Y		Y									Y	Y														Y												Y			
325	Y		Y									Y							Y			Y																		Y		Y
326	Y		Y									Y											Y						Y	Y												
327	Y		Y			Y		Y	Y			Y	Y			Y			Y				Y				Y	Y	Y	Y				Y		Y						
329	Y		Y			Y	Y	Y	Y	Y		Y	Y						Y			Y	Y	Y		Y	Y		Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
332	Y		Y		Y	Y		Y		Y		Y							Y				Y	Y			Y	Y	Y	Y						Y	Y			Y			Y
336	Y		Y		Y	Y	Y	Y	Y	Y		Y	Y			Y			Y		Y	Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
339	Y		Y		Y	Y	Y	Y	Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
340	Y		Y				Y		Y	Y		Y							Y			Y	Y			Y	Y		Y	Y					Y	Y	Y		Y	Y		Y
343	Y			Y				Y			Y	Y					Y			Y			Y																		
344	Y			Y				Y	Y		Y	Y		Y								Y	Y				Y								Y				Y	Y		Y
346	Y			Y		Y		Y			Y	Y			Y									Y			Y	Y				Y	Y						Y			
347	Y			Y	Y			Y			Y	Y				Y																						Y				
352	Y				Y	Y			Y			Y							Y				Y				Y	Y	Y	Y									Y			
354	Y				Y	Y	Y		Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y						Y		Y					Y
356	Y				Y				Y	Y		Y							Y					Y			Y								Y	Y			Y			
358	Y				Y	Y	Y	Y	Y	Y		Y				Y			Y			Y	Y	Y			Y		Y	Y	Y				Y	Y	Y				Y	Y
361	Y				Y					Y	Y	Y											Y					Y	Y	Y				Y	Y	Y				Y	
363	Y				Y	Y	Y			Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y									Y			Y		Y
365	Y				Y	Y					Y	Y							Y					Y				Y	Y								Y				Y
369	Y				Y	Y			Y	Y		Y							Y				Y	Y		Y	Y	Y	Y	Y					Y				Y	Y	Y	

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichornutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
370	Y				Y				Y	Y		Y	Y						Y								Y	Y	Y	Y						Y	Y		Y		
372	Y				Y	Y		Y	Y	Y		Y							Y				Y	Y				Y	Y	Y					Y		Y	Y				
373	Y				Y	Y	Y		Y	Y	Y	Y	Y						Y			Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
374	Y				Y	Y		Y	Y	Y		Y		Y					Y			Y	Y			Y	Y	Y	Y	Y					Y	Y	Y		Y	Y		Y
375	Y				Y					Y	Y	Y							Y				Y	Y				Y	Y	Y					Y	Y	Y		Y			Y
378	Y				Y				Y	Y		Y							Y					Y				Y									Y					
382	Y				Y	Y	Y	Y	Y	Y		Y	Y						Y	Y			Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	
383	Y				Y			Y				Y												Y											Y							
385	Y				Y				Y	Y		Y							Y					Y				Y	Y	Y					Y	Y	Y		Y			
386	Y				Y			Y				Y							Y	Y		Y	Y	Y			Y	Y	Y	Y					Y	Y	Y		Y	Y		Y
388	Y				Y			Y	Y	Y		Y							Y				Y			Y	Y	Y	Y	Y							Y		Y	Y		Y
395	Y				Y			Y		Y		Y							Y				Y	Y											Y	Y	Y					Y
398	Y				Y				Y	Y		Y				Y			Y				Y	Y		Y	Y	Y	Y	Y	Y					Y				Y	Y	Y
400	Y				Y	Y		Y	Y	Y		Y	Y						Y				Y	Y			Y	Y	Y	Y					Y	Y	Y		Y			Y
401	Y				Y	Y	Y	Y	Y	Y	Y	Y	Y						Y			Y	Y	Y			Y	Y	Y	Y	Y				Y	Y			Y	Y	Y	Y
402	Y				Y	Y			Y	Y		Y	Y						Y				Y	Y			Y	Y	Y	Y									Y			Y
404	Y				Y	Y	Y		Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y		Y
406	Y				Y				Y			Y	Y			Y						Y		Y			Y	Y	Y	Y					Y	Y	Y		Y		Y	Y
409	Y				Y	Y	Y	Y	Y	Y		Y	Y			Y						Y	Y	Y		Y	Y		Y	Y	Y				Y	Y	Y			Y	Y	Y
411	Y				Y	Y		Y	Y	Y		Y	Y						Y			Y		Y		Y	Y	Y	Y	Y						Y	Y		Y			Y
412	Y				Y			Y	Y	Y		Y							Y			Y	Y	Y		Y	Y	Y	Y	Y							Y	Y	Y			Y
416	Y				Y		Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
417	Y				Y			Y	Y	Y		Y							Y				Y				Y	Y	Y	Y					Y				Y			
421	Y				Y	Y			Y	Y		Y	Y	Y					Y			Y	Y	Y		Y	Y		Y	Y								Y				Y
422	Y				Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y		Y		Y	Y		Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
423	Y				Y	Y		Y	Y	Y		Y							Y				Y	Y		Y	Y	Y	Y	Y						Y	Y		Y			
424	Y				Y							Y										Y					Y											Y		Y		
427	Y				Y		Y		Y	Y		Y	Y						Y			Y	Y			Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
429	Y				Y		Y	Y	Y	Y	Y	Y	Y						Y			Y	Y				Y		Y	Y	Y				Y	Y	Y		Y		Y	Y
431	Y				Y	Y	Y		Y	Y		Y	Y						Y				Y	Y		Y	Y		Y	Y	Y				Y	Y			Y	Y	Y	
434	Y				Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
439	Y				Y		Y	Y	Y	Y	Y	Y	Y						Y		Y	Y	Y			Y	Y		Y	Y	Y		Y		Y	Y	Y	Y		Y	Y	Y
440	Y				Y	Y	Y	Y	Y	Y	Y	Y				Y			Y			Y	Y	Y		Y	Y	Y	Y	Y					Y	Y			Y		Y	Y
442	Y				Y	Y			Y	Y		Y										Y	Y			Y		Y	Y	Y						Y	Y				Y	Y
443	Y				Y			Y	Y			Y										Y					Y	Y														
445	Y				Y							Y																							Y							
446	Y				Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
448	Y				Y	Y		Y			Y	Y							Y				Y					Y		Y						Y	Y					Y
450	Y				Y	Y		Y	Y	Y		Y							Y			Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y			Y
453	Y				Y	Y	Y	Y	Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y	Y		Y		Y	Y	Y		Y	Y	Y	Y
454	Y				Y	Y			Y	Y	Y	Y		Y					Y				Y				Y	Y	Y	Y									Y		Y	
456	Y				Y	Y		Y	Y	Y	Y	Y		Y					Y				Y			Y	Y	Y	Y	Y		Y					Y	Y	Y			
460	Y				Y		Y			Y		Y	Y									Y	Y				Y		Y	Y	Y						Y		Y	Y		Y
464	Y				Y							Y	Y			Y																			Y							Y	
465	Y				Y	Y		Y				Y	Y									Y	Y	Y		Y		Y			Y				Y				Y		Y	
466	Y				Y	Y	Y	Y		Y		Y	Y						Y	Y		Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
471	Y				Y	Y	Y	Y	Y	Y		Y				Y			Y			Y	Y	Y				Y	Y	Y	Y				Y	Y	Y			Y		Y

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
472	Y				Y	Y	Y		Y	Y		Y	Y			Y			Y			Y	Y	Y			Y	Y	Y					Y	Y			Y	Y		Y
473	Y				Y	Y	Y		Y	Y		Y	Y						Y				Y	Y		Y	Y		Y					Y	Y	Y			Y		
476	Y				Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y				Y		Y		Y	Y	Y	Y
478	Y				Y	Y		Y	Y	Y		Y	Y						Y				Y	Y			Y	Y	Y					Y	Y	Y		Y			
480	Y				Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
481	Y				Y	Y		Y	Y	Y		Y							Y				Y	Y			Y	Y	Y	Y				Y	Y			Y	Y		
483	Y				Y				Y	Y		Y							Y				Y				Y	Y	Y							Y					
485	Y				Y	Y			Y	Y		Y							Y			Y	Y	Y		Y		Y	Y	Y					Y	Y	Y			Y		Y
486	Y				Y							Y													Y	Y															
487	Y				Y		Y					Y											Y						Y					Y						Y	
488	Y				Y			Y	Y	Y		Y	Y			Y			Y	Y		Y	Y	Y		Y	Y		Y	Y	Y		Y		Y	Y			Y	Y	Y
489	Y				Y	Y	Y	Y	Y	Y		Y	Y						Y			Y				Y	Y	Y	Y	Y	Y				Y	Y	Y		Y		Y	Y
491	Y				Y				Y	Y	Y	Y							Y				Y	Y				Y	Y	Y						Y					
492	Y				Y	Y		Y	Y	Y		Y		Y					Y				Y	Y			Y	Y	Y	Y					Y	Y	Y	Y	Y	Y		
493	Y				Y	Y		Y				Y							Y			Y	Y						Y								Y		Y			Y
494	Y				Y				Y			Y																	Y												
495	Y				Y	Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y			Y	Y	Y	Y
496	Y				Y					Y		Y							Y				Y					Y		Y											
497	Y				Y		Y		Y	Y		Y							Y					Y		Y		Y	Y	Y							Y					
499	Y				Y	Y	Y	Y	Y	Y		Y	Y			Y			Y	Y		Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
501	Y				Y	Y	Y	Y		Y	Y	Y	Y			Y			Y			Y		Y			Y	Y	Y	Y	Y				Y	Y			Y	Y	Y	Y
502	Y				Y	Y	Y	Y		Y		Y	Y						Y			Y	Y	Y			Y		Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
503	Y				Y	Y		Y	Y	Y		Y							Y				Y				Y	Y	Y	Y							Y		Y			
504	Y				Y		Y	Y	Y	Y		Y	Y									Y	Y	Y		Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
508	Y				Y					Y		Y							Y								Y														
509	Y				Y	Y	Y	Y	Y	Y	Y	Y	Y						Y				Y			Y	Y	Y	Y	Y	Y				Y	Y					Y	
513	Y				Y	Y	Y	Y	Y	Y	Y	Y	Y	Y		Y			Y			Y	Y	Y		Y	Y		Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
514	Y				Y	Y	Y	Y	Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
516	Y				Y				Y	Y	Y	Y				Y			Y				Y						Y	Y					Y	Y	Y					Y
517	Y				Y	Y		Y				Y										Y	Y																			
522	Y				Y	Y	Y	Y	Y	Y		Y	Y			Y			Y	Y	Y	Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
523	Y				Y					Y		Y							Y				Y	Y					Y	Y							Y					
524	Y				Y	Y	Y	Y	Y	Y		Y	Y						Y				Y	Y			Y		Y	Y						Y	Y		Y		Y	
527	Y				Y							Y																							Y						
529	Y				Y			Y	Y			Y				Y			Y				Y	Y			Y	Y	Y	Y					Y				Y	Y	Y	
530	Y				Y					Y		Y											Y	Y			Y		Y	Y									Y			Y
532	Y				Y			Y		Y		Y							Y				Y				Y		Y	Y							Y			Y		Y
533	Y				Y	Y	Y	Y	Y	Y	Y	Y	Y		Y	Y			Y	Y		Y		Y	Y	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y	Y
534	Y				Y	Y			Y	Y		Y	Y						Y		Y	Y							Y	Y							Y					
535	Y				Y	Y	Y	Y	Y			Y	Y			Y			Y	Y		Y		Y		Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
536	Y				Y	Y		Y	Y	Y		Y	Y									Y	Y	Y		Y		Y	Y	Y						Y	Y		Y	Y	Y	Y
537	Y				Y							Y	Y						Y							Y				Y	Y					Y			Y	Y		
538	Y				Y	Y	Y	Y	Y	Y		Y							Y			Y	Y	Y				Y	Y	Y	Y					Y	Y	Y		Y	Y		
539	Y				Y	Y	Y	Y	Y	Y	Y	Y	Y			Y			Y		Y	Y	Y	Y		Y	Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
541	Y				Y	Y	Y	Y	Y	Y	Y	Y	Y			Y			Y			Y	Y	Y			Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
544	Y				Y	Y	Y		Y	Y		Y	Y						Y				Y	Y		Y	Y	Y	Y	Y							Y	Y		Y	Y		
545	Y				Y	Y				Y		Y							Y									Y														

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
546	Y				Y					Y		Y							Y			Y	Y			Y	Y	Y							Y	Y			Y
547	Y				Y		Y	Y	Y	Y		Y	Y	Y		Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y		Y	Y	Y	Y
548	Y				Y					Y		Y											Y					Y								Y					
549	Y				Y	Y	Y		Y	Y		Y							Y				Y	Y		Y	Y	Y	Y	Y	Y					Y		Y	Y		
550	Y				Y	Y		Y	Y	Y		Y		Y					Y				Y	Y			Y	Y	Y	Y					Y	Y	Y		Y			
556	Y				Y	Y						Y							Y					Y														Y			
557	Y				Y	Y			Y			Y							Y																	Y					
562	Y				Y							Y							Y	Y		Y					Y								Y						
565	Y				Y	Y		Y	Y	Y		Y		Y					Y				Y				Y	Y	Y	Y	Y				Y	Y	Y	Y	Y	Y		
569	Y					Y				Y	Y	Y							Y				Y			Y		Y	Y	Y											
572	Y					Y						Y											Y	Y				Y													
575	Y					Y		Y	Y	Y		Y														Y	Y		Y	Y						Y					Y
577	Y					Y	Y	Y	Y	Y		Y	Y			Y	Y		Y			Y	Y	Y			Y	Y	Y	Y	Y		Y		Y	Y	Y				Y
582	Y					Y	Y					Y															Y	Y	Y	Y	Y										
585	Y					Y	Y	Y	Y			Y	Y			Y						Y	Y	Y		Y	Y	Y			Y				Y	Y			Y	Y	Y	Y
586	Y					Y						Y															Y														
588	Y					Y						Y								Y				Y														Y			
590	Y					Y	Y	Y	Y	Y		Y							Y			Y				Y		Y	Y	Y	Y				Y	Y				Y	Y	Y
592	Y					Y			Y			Y				Y							Y	Y				Y							Y	Y						
594	Y					Y						Y		Y							Y		Y	Y				Y												Y		
595	Y					Y	Y	Y	Y	Y		Y	Y						Y			Y	Y			Y	Y	Y	Y	Y	Y				Y	Y	Y			Y	Y	Y
596	Y					Y		Y			Y	Y								Y				Y			Y														
597	Y					Y		Y	Y	Y		Y		Y					Y					Y			Y	Y	Y	Y					Y	Y	Y	Y				
598	Y					Y					Y	Y					Y						Y				Y												Y	Y		
599	Y					Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y				Y	Y	Y					Y	Y	Y		Y	Y		Y
600	Y					Y						Y																								Y					
601	Y					Y	Y	Y	Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y					Y	Y	Y		Y			Y
602	Y					Y		Y				Y											Y	Y			Y	Y				Y					Y				
604	Y					Y			Y	Y		Y							Y				Y	Y		Y	Y	Y	Y	Y	Y					Y	Y		Y	Y		
605	Y					Y						Y								Y			Y					Y	Y	Y				Y	Y						
607	Y					Y						Y											Y	Y			Y	Y											Y			
609	Y					Y	Y	Y		Y		Y	Y						Y				Y	Y					Y	Y	Y					Y	Y			Y	Y	Y
611	Y					Y		Y		Y		Y							Y	Y				Y		Y	Y	Y	Y	Y	Y				Y		Y		Y	Y		
613	Y					Y		Y	Y	Y	Y	Y		Y			Y		Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y	Y	Y	Y		Y
616	Y					Y	Y	Y	Y	Y	Y	Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
617	Y					Y						Y		Y													Y	Y								Y						
620	Y					Y		Y	Y	Y	Y	Y							Y			Y		Y		Y			Y	Y	Y						Y		Y	Y	Y	Y
621	Y					Y		Y		Y		Y							Y				Y	Y			Y	Y		Y							Y		Y			
622	Y					Y		Y				Y											Y	Y			Y	Y	Y	Y									Y			
624	Y					Y					Y	Y							Y			Y	Y	Y				Y											Y		Y	
625	Y					Y		Y	Y	Y		Y				Y			Y				Y	Y	Y	Y	Y	Y	Y	Y					Y	Y	Y		Y			
629	Y					Y			Y	Y		Y	Y						Y	Y		Y		Y		Y	Y	Y	Y	Y	Y				Y	Y	Y	Y	Y	Y		Y
630	Y					Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y			Y		Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
631	Y					Y			Y	Y		Y			Y							Y	Y					Y	Y	Y	Y								Y	Y		Y
632	Y					Y						Y		Y								Y	Y	Y			Y	Y	Y	Y	Y								Y	Y	Y		Y
634	Y					Y				Y		Y											Y	Y		Y		Y														
636	Y					Y		Y				Y											Y	Y				Y	Y	Y					Y	Y			Y			

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
641	Y					Y		Y				Y											Y															Y	Y			
642	Y					Y		Y				Y												Y			Y	Y														
644	Y					Y		Y	Y	Y		Y		Y					Y	Y		Y	Y	Y		Y	Y	Y	Y	Y						Y			Y		Y
646	Y					Y		Y				Y		Y									Y	Y			Y	Y	Y	Y					Y	Y		Y	Y			
649	Y					Y		Y			Y	Y											Y				Y	Y										Y	Y			
650	Y					Y						Y																Y										Y				
652	Y					Y				Y		Y												Y				Y									Y		Y			
653	Y					Y		Y	Y			Y											Y	Y			Y										Y	Y			
654	Y					Y		Y			Y	Y											Y	Y			Y	Y														
656	Y					Y		Y	Y	Y		Y							Y				Y					Y	Y	Y	Y					Y							
657	Y					Y		Y			Y	Y											Y	Y		Y	Y											Y	Y			Y
658	Y					Y						Y											Y																			
660	Y					Y			Y	Y	Y	Y							Y			Y		Y			Y		Y	Y			Y				Y		Y			Y
661	Y					Y	Y	Y	Y	Y		Y							Y			Y		Y		Y		Y	Y					Y	Y			Y		Y		Y
665	Y					Y	Y	Y		Y		Y				Y						Y	Y	Y				Y	Y	Y					Y	Y			Y		Y	
666	Y					Y		Y	Y			Y							Y	Y			Y	Y		Y	Y	Y	Y									Y			
667	Y					Y						Y			Y				Y				Y					Y	Y	Y					Y	Y						Y
668	Y					Y	Y	Y	Y	Y		Y	Y		Y				Y			Y	Y			Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
669	Y					Y	Y	Y			Y	Y										Y	Y	Y			Y	Y		Y	Y		Y				Y		Y			Y
673	Y					Y				Y		Y							Y				Y			Y		Y			Y				Y					Y		
674	Y					Y						Y		Y		Y						Y	Y	Y		Y	Y	Y	Y	Y					Y	Y						
675	Y					Y		Y				Y		Y									Y	Y							Y		Y					Y	Y	Y			
676	Y					Y						Y												Y			Y	Y		Y			Y									
680	Y					Y			Y	Y		Y		Y					Y				Y	Y		Y			Y	Y						Y	Y					
682	Y					Y	Y	Y	Y	Y	Y	Y							Y			Y	Y	Y			Y			Y	Y					Y	Y	Y		Y	Y	Y
683	Y					Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y		Y	Y	Y	Y					Y	Y	Y			Y	Y
684	Y					Y		Y				Y																													
685	Y					Y				Y		Y							Y				Y	Y				Y							Y		Y		Y			
687	Y					Y		Y				Y		Y								Y	Y	Y			Y	Y										Y				
688	Y					Y					Y	Y							Y					Y		Y		Y							Y					Y	Y	Y
690	Y					Y	Y	Y	Y	Y		Y		Y		Y			Y				Y	Y				Y			Y							Y	Y		Y	Y
691	Y					Y		Y			Y	Y											Y	Y			Y	Y										Y	Y			
693	Y					Y		Y				Y											Y	Y			Y	Y										Y	Y			
697	Y					Y						Y											Y				Y	Y														
700	Y					Y				Y		Y							Y				Y				Y	Y									Y		Y			
701	Y					Y		Y		Y	Y	Y	Y						Y				Y	Y			Y	Y		Y						Y	Y		Y	Y		
702	Y					Y	Y	Y				Y										Y		Y		Y			Y	Y									Y	Y		
705	Y					Y						Y											Y	Y			Y	Y											Y			
707	Y					Y			Y	Y	Y	Y							Y	Y			Y	Y			Y	Y	Y	Y					Y	Y						Y
711	Y					Y						Y							Y				Y			Y		Y		Y							Y					
712	Y					Y		Y				Y		Y									Y				Y	Y			Y								Y	Y	Y		
715	Y					Y						Y																Y														
716	Y					Y		Y				Y		Y									Y	Y				Y									Y	Y			Y		
717	Y					Y						Y										Y				Y			Y		Y				Y					Y		Y
718	Y					Y						Y											Y							Y						Y	Y					
721	Y					Y	Y			Y			Y										Y	Y		Y		Y	Y	Y			Y					Y			Y	Y
723	Y					Y						Y				Y			Y			Y	Y					Y										Y				Y

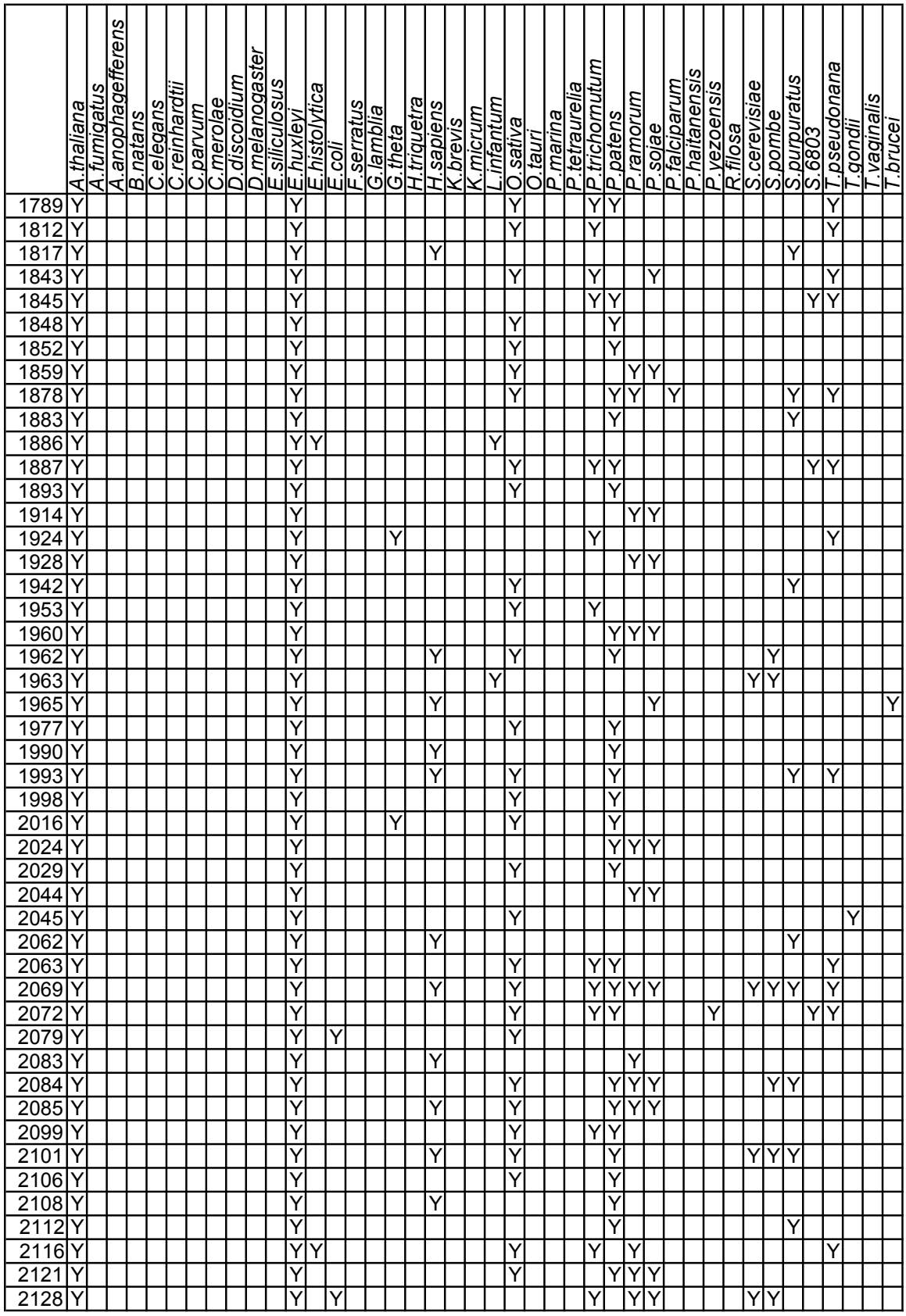
	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
724	Y					Y		Y	Y	Y		Y	Y									Y			Y	Y	Y	Y	Y					Y	Y	Y					
725	Y					Y	Y					Y							Y	Y		Y	Y	Y		Y	Y		Y	Y								Y	Y		
726	Y					Y		Y				Y												Y			Y	Y										Y				
728	Y					Y	Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y			Y
731	Y					Y	Y		Y	Y		Y							Y					Y			Y		Y	Y	Y						Y		Y	Y	
732	Y					Y			Y	Y		Y							Y								Y	Y	Y								Y		Y		
733	Y					Y	Y	Y	Y		Y	Y	Y											Y		Y	Y		Y	Y	Y				Y	Y	Y		Y	Y		Y
738	Y					Y		Y			Y	Y											Y	Y			Y	Y				Y						Y	Y			
739	Y					Y		Y	Y	Y		Y							Y			Y		Y		Y	Y	Y	Y	Y				Y	Y	Y	Y	Y			
742	Y					Y	Y	Y	Y	Y		Y							Y			Y	Y	Y			Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
743	Y					Y	Y		Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y			Y	Y	
747	Y					Y		Y	Y	Y	Y	Y							Y				Y			Y	Y	Y	Y	Y					Y	Y			Y	Y	
749	Y					Y			Y	Y		Y							Y			Y	Y	Y				Y	Y	Y							Y		Y		
750	Y					Y			Y	Y		Y	Y						Y	Y			Y	Y			Y	Y	Y		Y				Y	Y	Y			Y	Y	Y
751	Y					Y		Y	Y	Y	Y	Y							Y				Y	Y			Y	Y	Y	Y					Y	Y	Y	Y	Y			
752	Y					Y				Y		Y												Y				Y														
753	Y					Y	Y					Y										Y	Y	Y		Y	Y	Y	Y	Y									Y			
758	Y					Y			Y	Y		Y							Y				Y					Y	Y	Y							Y					
760	Y					Y					Y	Y											Y	Y			Y	Y														
761	Y					Y						Y								Y						Y			Y	Y							Y		Y	Y	Y	
762	Y					Y	Y	Y	Y	Y		Y	Y						Y			Y	Y	Y			Y	Y	Y	Y					Y	Y	Y		Y	Y		Y
763	Y					Y		Y				Y		Y							Y		Y	Y			Y		Y	Y					Y	Y		Y	Y			
765	Y					Y		Y	Y	Y		Y		Y					Y				Y	Y			Y	Y	Y	Y	Y				Y	Y		Y	Y		Y	
767	Y					Y			Y	Y		Y	Y						Y				Y	Y		Y	Y	Y	Y	Y	Y					Y	Y		Y		Y	
768	Y					Y						Y																Y	Y							Y					
769	Y					Y		Y	Y	Y		Y							Y				Y	Y		Y	Y	Y								Y	Y		Y	Y		
771	Y					Y						Y								Y			Y	Y			Y	Y										Y				
773	Y					Y						Y								Y			Y				Y	Y								Y			Y			
776	Y					Y	Y	Y	Y			Y	Y			Y			Y			Y	Y	Y	Y	Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
778	Y					Y			Y		Y	Y											Y				Y		Y	Y									Y			
779	Y					Y						Y	Y							Y										Y							Y					
781	Y					Y						Y										Y					Y	Y										Y				
782	Y					Y					Y	Y											Y				Y	Y		Y									Y			
783	Y					Y		Y		Y		Y	Y									Y		Y		Y		Y		Y									Y	Y	Y	
784	Y					Y						Y		Y									Y	Y			Y	Y	Y	Y						Y	Y		Y			
786	Y					Y						Y											Y				Y	Y											Y			
787	Y					Y			Y	Y		Y				Y			Y			Y	Y	Y		Y		Y									Y				Y	Y
789	Y					Y		Y		Y	Y	Y	Y			Y	Y	Y	Y	Y	Y	Y		Y	Y	Y			Y	Y	Y						Y		Y	Y		Y
790	Y					Y	Y	Y		Y		Y																Y	Y		Y				Y					Y		
791	Y					Y						Y																Y													
795	Y					Y						Y											Y				Y	Y													
796	Y					Y	Y	Y				Y		Y						Y			Y				Y	Y											Y	Y		
798	Y					Y		Y	Y			Y		Y									Y	Y			Y	Y	Y	Y									Y	Y		
799	Y					Y	Y		Y	Y		Y							Y			Y	Y					Y	Y	Y							Y			Y		Y
802	Y					Y						Y		Y									Y	Y							Y							Y		Y		
803	Y					Y	Y	Y				Y		Y									Y	Y			Y	Y			Y		Y						Y	Y		
804	Y					Y			Y		Y	Y										Y	Y					Y	Y	Y											Y

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
805	Y					Y		Y				Y																												
806	Y					Y	Y		Y			Y							Y							Y		Y	Y							Y		Y		Y	
807	Y					Y		Y			Y	Y		Y								Y	Y			Y	Y		Y									Y			
809	Y					Y		Y	Y	Y		Y	Y						Y			Y			Y			Y	Y					Y	Y	Y		Y	Y	Y	Y
810	Y					Y	Y		Y			Y	Y						Y	Y		Y	Y	Y		Y	Y	Y	Y						Y	Y		Y		Y	Y
814	Y					Y		Y			Y	Y					Y					Y	Y			Y					Y						Y	Y			
815	Y					Y						Y							Y			Y	Y			Y	Y	Y	Y	Y					Y	Y	Y		Y		Y
816	Y					Y		Y			Y	Y		Y		Y			Y	Y			Y	Y			Y		Y	Y					Y	Y	Y	Y	Y		Y	
817	Y					Y		Y	Y			Y	Y																					Y							
818	Y					Y						Y											Y					Y										Y			
819	Y					Y	Y		Y	Y		Y							Y				Y				Y	Y	Y	Y	Y					Y	Y		Y	Y		
821	Y					Y	Y	Y	Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	Y
824	Y					Y						Y		Y									Y					Y										Y			
828	Y					Y				Y		Y							Y																	Y					
830	Y					Y						Y											Y	Y			Y					Y						Y	Y			
833	Y					Y			Y	Y		Y							Y			Y	Y	Y			Y	Y	Y	Y					Y	Y						Y
834	Y					Y						Y											Y				Y											Y			
835	Y						Y	Y	Y	Y		Y	Y									Y	Y		Y	Y	Y	Y	Y	Y					Y	Y	Y		Y	Y		Y
838	Y						Y		Y			Y		Y		Y						Y		Y	Y	Y				Y									Y	Y		Y
846	Y						Y	Y	Y	Y		Y							Y				Y	Y		Y		Y	Y	Y	Y						Y			Y		
849	Y						Y	Y				Y				Y			Y																Y	Y						
856	Y						Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
858	Y						Y	Y				Y	Y							Y							Y								Y				Y	Y		
859	Y						Y					Y																							Y							
860	Y						Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
862	Y						Y		Y	Y		Y							Y			Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
864	Y						Y		Y			Y												Y		Y			Y	Y												
865	Y						Y	Y	Y			Y	Y										Y	Y			Y	Y	Y		Y				Y	Y	Y		Y	Y	Y	Y
873	Y						Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
874	Y						Y			Y		Y							Y			Y	Y	Y			Y			Y								Y			Y
877	Y						Y					Y																	Y										Y		
878	Y						Y	Y	Y	Y	Y	Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y					Y	Y			Y	Y	Y	Y
879	Y						Y	Y	Y	Y		Y	Y						Y								Y		Y	Y	Y						Y	Y			Y		
880	Y						Y	Y	Y	Y		Y	Y			Y			Y	Y		Y	Y	Y					Y	Y	Y				Y		Y		Y	Y	Y	Y
881	Y						Y					Y															Y	Y		Y					Y					Y		
883	Y						Y		Y	Y		Y	Y						Y				Y	Y		Y	Y		Y	Y	Y				Y	Y			Y		Y	Y
895	Y						Y	Y				Y										Y					Y		Y	Y	Y		Y			Y		Y	Y	Y	
897	Y						Y	Y	Y			Y	Y			Y			Y				Y	Y					Y	Y												
899	Y						Y	Y	Y	Y		Y							Y				Y				Y	Y	Y	Y	Y							Y		Y	Y		
900	Y						Y					Y															Y									Y						
905	Y						Y	Y	Y	Y		Y	Y			Y						Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y			Y	Y	Y	Y
908	Y						Y	Y	Y			Y																			Y										
913	Y						Y		Y	Y		Y			Y				Y				Y			Y	Y	Y	Y	Y							Y		Y	Y		
917	Y						Y					Y											Y						Y	Y	Y								Y	Y	
921	Y						Y		Y			Y										Y																			Y
926	Y						Y		Y	Y	Y	Y				Y			Y			Y	Y	Y				Y		Y	Y				Y	Y	Y			Y	Y	Y
927	Y						Y		Y			Y							Y				Y	Y		Y	Y	Y	Y	Y	Y					Y	Y			Y	Y	Y

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
928	Y						Y		Y	Y		Y	Y						Y				Y			Y		Y	Y	Y					Y				Y			
929	Y						Y	Y	Y	Y		Y	Y			Y			Y				Y	Y		Y		Y	Y	Y	Y				Y	Y	Y			Y	Y	Y
936	Y						Y	Y	Y			Y	Y						Y				Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y	Y		
938	Y						Y	Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y		Y	Y	Y				Y	Y			Y	Y	Y	Y
951	Y						Y		Y	Y		Y							Y	Y			Y			Y		Y			Y									Y	Y	
953	Y						Y	Y	Y	Y	Y	Y		Y					Y	Y	Y	Y				Y	Y		Y	Y	Y		Y		Y	Y	Y	Y	Y		Y
954	Y						Y	Y	Y		Y	Y	Y			Y						Y	Y	Y		Y	Y	Y	Y	Y	Y				Y				Y	Y	Y	Y
959	Y							Y				Y											Y	Y			Y	Y									Y				
960	Y							Y				Y																Y										Y	Y		
961	Y							Y	Y	Y		Y							Y				Y	Y			Y		Y	Y					Y	Y	Y	Y			Y	
962	Y							Y				Y												Y			Y	Y										Y			
965	Y							Y				Y											Y	Y				Y	Y	Y	Y					Y			Y	Y		
966	Y							Y	Y	Y		Y		Y					Y	Y			Y	Y		Y	Y	Y			Y					Y	Y	Y	Y	Y		
968	Y							Y				Y												Y		Y	Y				Y							Y	Y	Y		
970	Y							Y				Y											Y	Y														Y				
971	Y							Y	Y	Y	Y	Y							Y	Y			Y	Y	Y	Y	Y	Y	Y	Y	Y		Y		Y	Y	Y		Y	Y		Y
973	Y							Y			Y	Y		Y				Y	Y				Y	Y		Y	Y		Y		Y			Y			Y	Y	Y	Y		
975	Y							Y				Y											Y	Y																		
976	Y							Y				Y							Y								Y								Y			Y				
980	Y							Y				Y		Y									Y	Y														Y		Y		
986	Y							Y				Y											Y				Y	Y										Y	Y			
989	Y							Y	Y		Y	Y										Y	Y	Y			Y								Y	Y				Y			Y
991	Y							Y	Y	Y		Y											Y							Y	Y											
992	Y							Y	Y			Y	Y						Y				Y	Y		Y	Y	Y	Y	Y							Y		Y			
994	Y							Y			Y	Y	Y	Y									Y	Y		Y	Y	Y	Y	Y					Y			Y	Y			
995	Y							Y				Y											Y	Y				Y											Y			
996	Y							Y				Y																			Y									Y		
998	Y							Y				Y		Y									Y	Y			Y	Y			Y							Y	Y	Y		
1000	Y							Y	Y			Y		Y									Y	Y			Y	Y			Y							Y	Y	Y		
1008	Y							Y			Y	Y												Y			Y											Y	Y			
1012	Y							Y	Y	Y	Y	Y			Y				Y					Y			Y	Y	Y	Y		Y				Y	Y			Y	Y		Y
1018	Y							Y	Y	Y	Y	Y										Y	Y				Y								Y	Y			Y			Y
1020	Y							Y				Y								Y			Y	Y			Y												Y				
1023	Y							Y		Y		Y	Y						Y																		Y					Y
1031	Y							Y	Y			Y							Y			Y	Y	Y		Y	Y	Y							Y	Y	Y					Y
1034	Y							Y				Y																Y														
1037	Y							Y				Y	Y						Y							Y			Y							Y							
1039	Y							Y				Y											Y	Y		Y	Y	Y	Y	Y	Y							Y		Y		
1044	Y							Y	Y	Y	Y	Y		Y					Y				Y	Y			Y	Y	Y	Y					Y	Y	Y	Y	Y	Y		
1046	Y							Y	Y			Y												Y			Y									Y		Y				
1048	Y							Y	Y	Y	Y	Y	Y	Y					Y				Y	Y		Y	Y											Y	Y	Y	Y		
1055	Y							Y	Y	Y	Y	Y			Y				Y			Y	Y	Y		Y			Y	Y	Y					Y	Y	Y			Y	Y	Y
1072	Y							Y				Y	Y		Y							Y					Y			Y									Y	Y	Y	Y
1075	Y							Y				Y											Y				Y	Y										Y	Y			
1080	Y							Y	Y	Y		Y		Y					Y				Y	Y		Y	Y	Y	Y	Y	Y				Y	Y	Y	Y	Y	Y		
1082	Y							Y	Y	Y	Y	Y		Y					Y	Y	Y	Y	Y	Y		Y	Y	Y	Y	Y	Y		Y		Y	Y	Y		Y	Y		
1083	Y							Y	Y			Y							Y										Y										Y	Y		

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
1088	Y							Y				Y											Y	Y			Y	Y				Y						Y	Y			
1093	Y							Y				Y												Y			Y	Y									Y	Y			
1095	Y							Y				Y											Y	Y			Y	Y		Y				Y	Y						
1101	Y							Y				Y		Y					Y				Y				Y	Y	Y	Y	Y					Y	Y	Y	Y	Y		
1107	Y							Y	Y	Y		Y							Y			Y					Y	Y	Y	Y								Y			Y
1119	Y							Y	Y	Y		Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y			Y	Y	Y	Y	Y	Y	Y	Y
1120	Y							Y	Y	Y	Y	Y					Y		Y			Y	Y	Y		Y	Y		Y	Y				Y	Y	Y		Y			Y
1131	Y							Y	Y	Y		Y				Y						Y	Y	Y			Y	Y	Y	Y	Y					Y	Y		Y	Y	Y	Y
1132	Y							Y	Y			Y		Y					Y				Y	Y		Y	Y	Y									Y	Y	Y			
1134	Y							Y				Y											Y	Y			Y	Y											Y			
1143	Y							Y				Y								Y																					
1146	Y							Y				Y							Y				Y	Y		Y	Y	Y	Y					Y		Y	Y	Y		Y	Y
1150	Y							Y				Y											Y	Y			Y	Y										Y			
1152	Y							Y				Y											Y	Y					Y	Y											
1162	Y							Y	Y			Y		Y									Y	Y				Y													
1175	Y							Y				Y											Y	Y			Y	Y											Y			
1185	Y							Y	Y			Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y						Y		Y			Y
1190	Y							Y	Y			Y		Y					Y					Y		Y		Y	Y	Y				Y	Y		Y	Y			
1193	Y								Y	Y		Y											Y				Y	Y	Y	Y								Y			
1198	Y								Y	Y		Y							Y			Y		Y		Y	Y	Y	Y	Y								Y		Y	Y
1200	Y								Y	Y		Y							Y				Y	Y			Y	Y	Y	Y				Y	Y			Y	Y	Y	
1220	Y								Y	Y		Y				Y			Y				Y			Y	Y		Y	Y						Y					
1230	Y								Y	Y		Y															Y	Y	Y	Y											
1232	Y								Y			Y											Y	Y				Y						Y							
1234	Y								Y	Y		Y				Y			Y			Y	Y	Y		Y	Y	Y	Y	Y				Y	Y			Y	Y	Y	Y
1235	Y								Y	Y		Y															Y							Y				Y		Y	
1240	Y								Y		Y	Y		Y								Y		Y			Y		Y	Y						Y		Y			
1244	Y								Y	Y		Y	Y						Y				Y		Y	Y	Y		Y	Y				Y	Y	Y		Y		Y	Y
1263	Y								Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y	Y				Y	Y		Y	Y		Y
1267	Y								Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y	Y			Y	Y	Y					Y
1278	Y								Y			Y							Y				Y			Y		Y													Y
1280	Y								Y	Y		Y	Y			Y			Y	Y			Y				Y	Y	Y	Y						Y		Y	Y	Y	
1294	Y								Y			Y							Y				Y																		
1299	Y								Y	Y		Y				Y			Y			Y	Y	Y		Y	Y	Y	Y	Y	Y			Y	Y	Y		Y	Y	Y	Y
1305	Y								Y			Y							Y					Y			Y	Y								Y					
1323	Y								Y		Y	Y	Y			Y			Y			Y	Y	Y		Y	Y	Y	Y	Y				Y	Y			Y	Y	Y	
1324	Y								Y		Y	Y		Y					Y			Y		Y			Y	Y						Y		Y	Y	Y			
1325	Y								Y	Y		Y	Y						Y				Y	Y		Y	Y	Y	Y	Y						Y	Y		Y			Y
1329	Y								Y	Y		Y	Y	Y					Y				Y					Y													Y	
1333	Y								Y			Y	Y						Y				Y	Y			Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	
1341	Y								Y			Y		Y								Y	Y				Y	Y	Y	Y				Y					Y		Y
1344	Y								Y	Y		Y							Y			Y	Y	Y		Y	Y	Y	Y	Y						Y		Y	Y		Y
1350	Y								Y			Y											Y				Y											Y	Y			
1376	Y								Y			Y										Y				Y	Y														
1389	Y								Y	Y		Y	Y			Y			Y			Y	Y	Y			Y	Y	Y	Y	Y				Y	Y	Y		Y	Y	Y	Y
1392	Y								Y	Y		Y							Y			Y	Y	Y		Y		Y	Y	Y				Y		Y			Y	Y	Y
1397	Y								Y			Y				Y						Y		Y			Y		Y	Y					Y			Y		Y	

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
1419	Y								Y	Y		Y							Y			Y	Y	Y			Y	Y						Y	Y	Y					Y
1429	Y								Y	Y		Y							Y				Y	Y		Y	Y	Y	Y	Y								Y			
1441	Y								Y	Y		Y	Y						Y			Y	Y	Y		Y	Y	Y	Y	Y					Y	Y	Y		Y	Y	Y	
1451	Y									Y		Y				Y												Y													
1454	Y									Y		Y	Y						Y			Y	Y	Y		Y		Y	Y	Y	Y					Y	Y			Y		Y
1463	Y									Y		Y																													Y
1469	Y									Y		Y							Y									Y	Y									Y			
1472	Y									Y		Y							Y					Y			Y										Y			Y		
1476	Y									Y		Y										Y		Y		Y									Y	Y		Y				Y
1478	Y									Y		Y							Y										Y	Y	Y										
1480	Y									Y		Y											Y	Y											Y						Y	Y
1487	Y									Y		Y							Y			Y	Y	Y			Y	Y	Y	Y					Y	Y				Y	
1493	Y									Y		Y											Y							Y									Y		
1502	Y									Y	Y	Y							Y				Y					Y													
1503	Y									Y		Y	Y						Y			Y	Y	Y		Y			Y	Y	Y				Y	Y			Y	Y	Y	Y
1506	Y									Y		Y							Y								Y	Y	Y								Y				
1507	Y									Y		Y							Y				Y					Y							Y	Y						
1532	Y									Y	Y	Y	Y	Y									Y			Y	Y	Y	Y	Y							Y		Y	Y	Y	
1544	Y									Y		Y							Y				Y					Y													
1547	Y									Y		Y							Y				Y												Y	Y						
1548	Y									Y		Y							Y				Y					Y	Y		Y					Y	Y		Y	Y		
1558	Y									Y		Y							Y									Y													
1559	Y									Y		Y							Y				Y			Y	Y	Y	Y	Y					Y	Y	Y		Y			
1562	Y									Y		Y	Y																Y	Y					Y	Y						
1577	Y									Y		Y							Y									Y							Y				Y			Y
1579	Y									Y		Y	Y						Y				Y														Y		Y			
1590	Y									Y		Y											Y			Y	Y	Y							Y				Y	Y		
1597	Y									Y		Y							Y					Y												Y			Y			
1618	Y									Y		Y											Y																		
1628	Y									Y		Y							Y								Y			Y					Y	Y					
1630	Y									Y		Y							Y			Y							Y	Y											
1632	Y									Y		Y							Y																						
1636	Y									Y		Y																	Y	Y					Y	Y		Y			
1653	Y									Y		Y							Y	Y														Y		Y		Y			
1655	Y									Y		Y											Y					Y												Y	Y
1667	Y									Y		Y	Y	Y													Y			Y				Y	Y	Y		Y		Y	Y
1675	Y									Y		Y							Y								Y							Y	Y			Y			Y
1678	Y									Y		Y		Y					Y				Y					Y	Y	Y					Y	Y	Y	Y	Y			
1679	Y									Y		Y												Y					Y	Y					Y	Y						
1686	Y									Y		Y							Y			Y	Y	Y		Y	Y	Y	Y	Y					Y	Y	Y		Y		Y	Y
1687	Y									Y	Y	Y							Y				Y	Y			Y	Y		Y									Y			Y
1691	Y									Y		Y							Y				Y	Y					Y	Y					Y	Y						
1695	Y										Y	Y		Y													Y												Y			
1708	Y										Y	Y														Y	Y	Y											Y	Y		
1715	Y										Y	Y																			Y								Y	Y		
1739	Y										Y	Y		Y																										Y		
1778	Y											Y											Y					Y	Y	Y							Y		Y			


A.thaliana
A.fumigatus
A.anophagefferens
B.natans
C.elegans
C.reinhardtii
C.parvum
C.merolae
D.discoidium
D.melanogaster
E.siliculosus
E.huxleyi
E.histolytica
E.coli
F.serratus
G.lamblia
G.theta
H.triquetra
H.sapiens
K.brevis
K.micrum
L.infantum
O.sativa
O.tauri
P.marina
P.tetraurelia
P.trichomutum
P.patens
P.ramorum
P.sojae
P.falciparum
P.haitanensis
P.yezoensis
R.filosa
S.cerevisiae
S.pombe
S.purpuratus
S.6803
T.pseudonana
T.gondii
T.vaginalis
T.brucei
1789
1812
1817
1843
1845
1848
1852
1859
1878
1883
1886
1887
1893
1914
1924
1928
1942
1953
1960
1962
1963
1965
1977
1990
1993
1998
2016
2024
2029
2044
2045
2062
2063
2069
2072
2079
2083
2084
2085
2099
2101
2106
2108
2112
2116
2121
2128

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
2160	Y											Y							Y				Y					Y													
2164	Y											Y											Y					Y										Y				
2175	Y											Y		Y									Y				Y	Y	Y	Y						Y						
2176	Y											Y																Y	Y	Y												
2181	Y											Y											Y				Y	Y	Y													
2186	Y											Y							Y				Y					Y	Y	Y							Y						
2207	Y											Y											Y															Y				
2210	Y											Y											Y												Y	Y	Y		Y		Y	
2218	Y											Y											Y					Y									Y					
2226	Y											Y											Y					Y														
2241	Y											Y											Y				Y								Y	Y						
2250	Y											Y											Y				Y								Y		Y					Y
2251	Y											Y							Y				Y				Y	Y									Y		Y			
2265	Y											Y										Y	Y				Y	Y												Y		Y
2276	Y											Y							Y																Y	Y				Y	
2304	Y											Y											Y				Y	Y													
2312	Y											Y		Y									Y					Y									Y					
2313	Y											Y							Y				Y					Y								Y			Y		
2321	Y											Y										Y						Y	Y	Y					Y		Y				
2323	Y											Y																Y	Y												
2338		Y	Y									Y															Y											Y			
2381		Y			Y				Y	Y		Y			Y					Y																		Y			
2384		Y			Y							Y							Y			Y						Y	Y					Y		Y	Y	Y		Y
2396		Y			Y		Y	Y				Y				Y									Y			Y	Y	Y		Y		Y			Y	Y		
2403		Y			Y		Y	Y	Y	Y		Y	Y		Y				Y			Y	Y			Y	Y	Y	Y	Y				Y	Y	Y	Y	Y		Y
2509		Y				Y			Y			Y					Y					Y						Y	Y						Y					Y
2540		Y				Y						Y																				Y								
2579		Y				Y						Y											Y				Y	Y												
2604		Y				Y			Y			Y													Y				Y									Y		
2660		Y					Y				Y	Y																							Y				Y	
3601			Y									Y											Y				Y		Y									Y		
3602			Y									Y	Y									Y																		
3605			Y									Y																							Y					Y
3611			Y									Y															Y		Y	Y							Y		
3616			Y									Y																	Y	Y									
3618			Y									Y							Y																	Y	Y			
3620			Y									Y	Y			Y			Y			Y					Y	Y		Y					Y		Y	Y		
3621			Y									Y		Y													Y										Y			
3622			Y									Y							Y								Y		Y	Y						Y		Y			
3623			Y									Y							Y									Y												
3624			Y									Y																	Y						Y					
3625			Y									Y							Y																	Y				
3627			Y									Y																Y	Y											
3630			Y									Y					Y																							Y
3637			Y									Y															Y										Y			
3638			Y									Y															Y										Y			
3640			Y									Y							Y								Y								Y	Y			Y	Y	

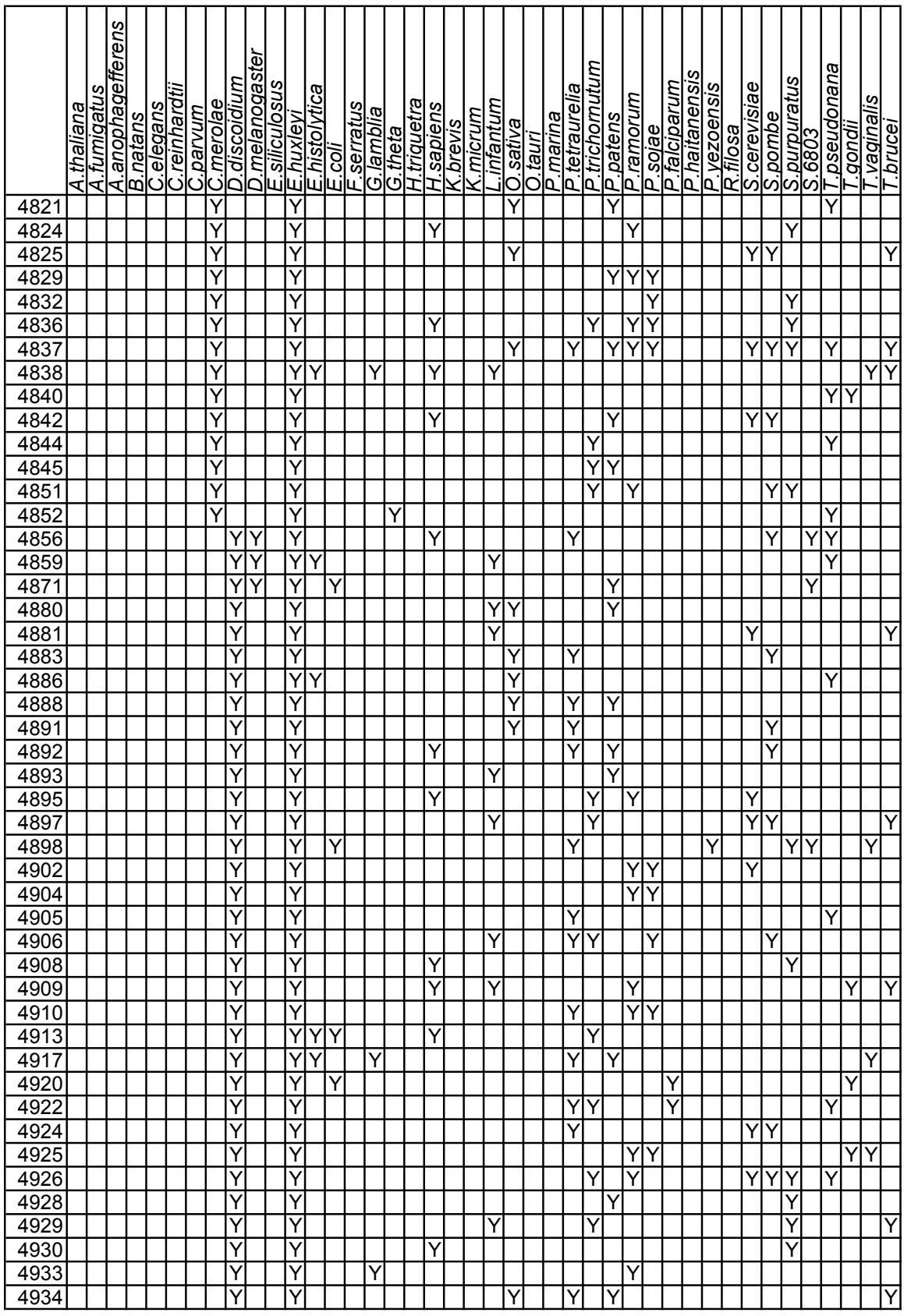
	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichomutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
3641			Y									Y	Y																									Y			
3642			Y									Y											Y				Y		Y	Y												
3643			Y									Y															Y											Y				
3648			Y									Y															Y											Y				
3650			Y									Y						Y					Y					Y	Y									Y				
3653			Y									Y															Y											Y				
3654			Y									Y											Y			Y		Y													Y	
3656			Y									Y															Y											Y				
3658			Y									Y																	Y							Y						
3664			Y									Y																	Y	Y												Y
3667			Y									Y																	Y	Y								Y	Y		
3669			Y									Y															Y			Y								Y			Y
3673			Y									Y											Y				Y											Y				
3674			Y									Y											Y					Y														
3678			Y									Y	Y										Y															Y				
3685			Y									Y															Y									Y		Y				
3688			Y									Y															Y											Y				
3691			Y									Y															Y		Y	Y												
3694			Y									Y											Y				Y		Y													
3697			Y									Y															Y		Y													
3710				Y								Y			Y												Y											Y				
3711				Y								Y																	Y												Y	
3730					Y	Y						Y							Y				Y				Y	Y	Y	Y						Y	Y					
3743					Y	Y						Y																	Y					Y							
3760					Y	Y						Y							Y			Y	Y				Y										Y				
3765					Y	Y						Y							Y																						
3780					Y		Y					Y	Y																		Y										
3816					Y			Y	Y	Y		Y		Y					Y			Y	Y			Y		Y							Y	Y	Y				Y	
3819					Y			Y				Y			Y											Y			Y									Y				Y
3830					Y			Y				Y																										Y			
3853					Y			Y	Y			Y							Y																	Y	Y				
3876					Y				Y			Y							Y			Y	Y														Y				
3877					Y				Y	Y		Y							Y											Y				Y		Y			Y		Y
3904					Y				Y			Y							Y																	Y					
3913					Y				Y			Y															Y		Y						Y	Y					
3924					Y					Y		Y																						Y	Y	Y					
3930					Y					Y		Y															Y		Y												
3932					Y					Y		Y																Y			Y					Y					
3963					Y					Y		Y			Y				Y			Y					Y	Y	Y											Y	Y	Y
3968					Y					Y	Y	Y							Y																Y	Y					
3973					Y					Y		Y																				Y						Y			
3986					Y					Y		Y															Y														
3992					Y						Y	Y	Y				Y									Y	Y									Y	Y		Y		Y	Y
3997					Y						Y	Y		Y					Y			Y					Y		Y	Y						Y		Y			
3999					Y						Y	Y														Y	Y		Y	Y											Y
4004					Y						Y	Y														Y									Y		Y		Y		
4012					Y							Y																	Y						Y						

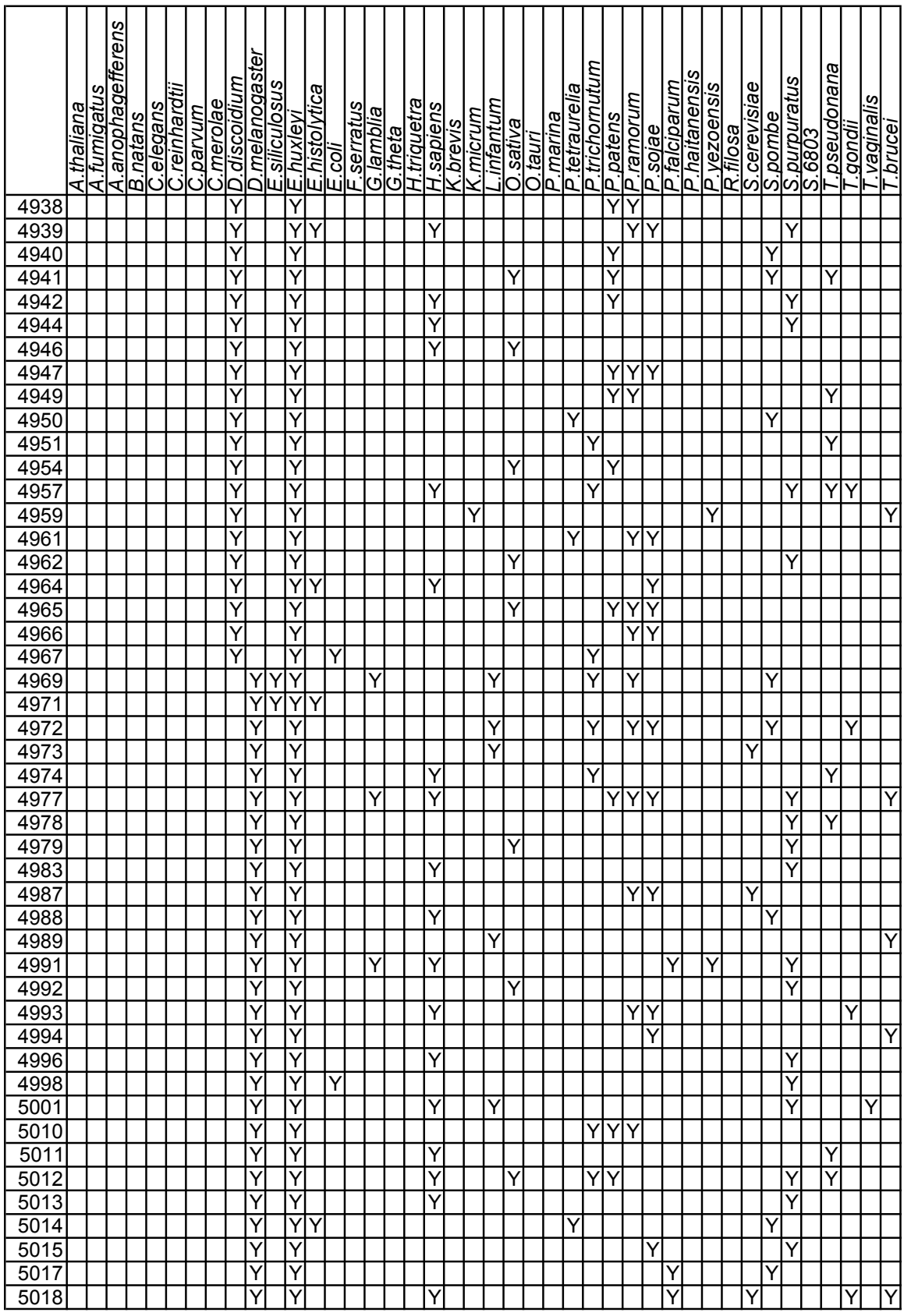
	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichornutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
4013					Y							Y								Y										Y						Y				
4019					Y							Y							Y							Y			Y	Y					Y	Y					
4020					Y							Y							Y										Y	Y						Y					
4021					Y							Y								Y																		Y			
4023					Y							Y											Y				Y				Y							Y	Y		
4031					Y							Y										Y													Y	Y					Y
4035					Y							Y							Y								Y	Y	Y							Y				Y	Y
4038					Y							Y				Y																		Y		Y		Y			
4040					Y							Y							Y																	Y					
4042					Y							Y										Y							Y												Y
4043					Y							Y										Y				Y			Y	Y											Y
4052					Y							Y											Y				Y														
4053					Y							Y											Y				Y			Y											
4054					Y							Y							Y										Y	Y											
4055					Y							Y							Y										Y	Y							Y		Y			
4056					Y							Y											Y															Y			
4058					Y							Y																	Y	Y											Y
4066					Y							Y															Y								Y	Y					
4067					Y							Y							Y							Y	Y	Y		Y								Y	Y		
4068					Y							Y		Y								Y	Y				Y	Y	Y	Y									Y			
4069					Y							Y																Y	Y	Y											
4071					Y							Y																Y	Y	Y							Y				
4072					Y							Y																	Y	Y											
4076					Y							Y	Y			Y			Y			Y	Y				Y	Y		Y					Y	Y			Y		Y	Y
4077					Y							Y							Y								Y								Y		Y					
4078					Y							Y							Y			Y																			
4082					Y							Y		Y																												Y
4083					Y							Y							Y								Y											Y			
4086					Y							Y	Y			Y			Y			Y				Y		Y	Y							Y		Y	Y		
4087					Y							Y							Y																	Y					
4092					Y							Y							Y							Y			Y	Y							Y				
4094					Y							Y															Y											Y			
4096					Y							Y	Y						Y									Y	Y	Y		Y		Y	Y	Y					
4099					Y							Y														Y													Y		
4101					Y							Y																Y										Y			
4103					Y							Y							Y								Y	Y	Y						Y						
4111					Y							Y				Y										Y														Y	
4116					Y							Y							Y															Y	Y						
4118					Y							Y										Y	Y					Y									Y	Y			Y
4121					Y							Y																		Y							Y	Y		Y
4122					Y							Y																									Y			Y
4125					Y							Y														Y													Y	
4126					Y							Y																						Y	Y					
4128					Y							Y				Y							Y						Y	Y							Y				
4129					Y							Y											Y				Y														
4130					Y							Y							Y																		Y				
4133					Y							Y																						Y	Y					Y

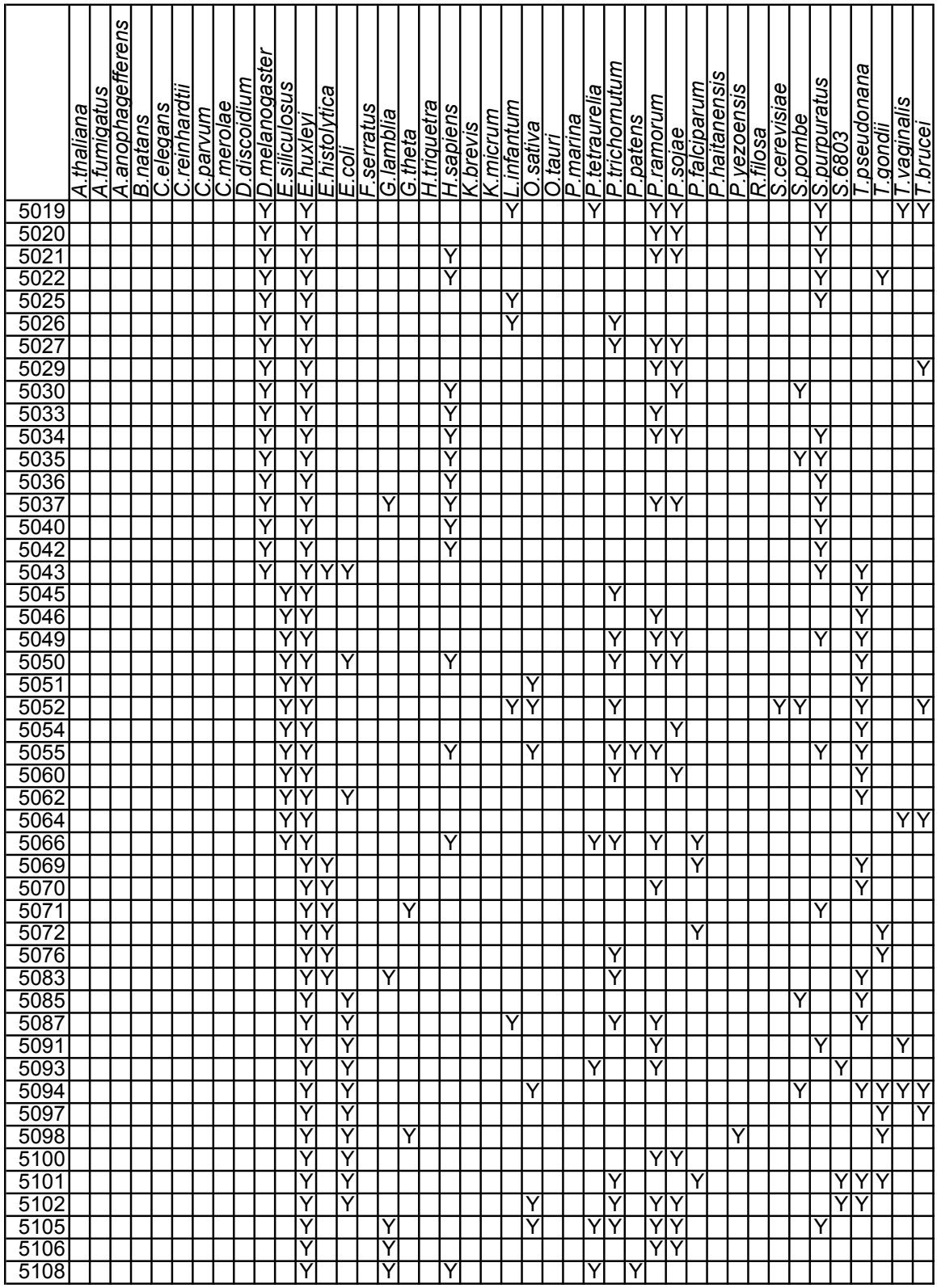
	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichornutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
4134					Y							Y	Y																												Y
4139					Y							Y																									Y					
4140					Y							Y																Y	Y												
4143					Y							Y							Y			Y			Y			Y							Y	Y		Y	Y		Y
4146					Y							Y											Y					Y	Y									Y			
4147					Y							Y							Y							Y												Y			
4149					Y							Y		Y															Y	Y						Y					
4150					Y							Y																	Y	Y						Y					Y
4152					Y							Y															Y									Y					
4155					Y							Y											Y				Y														
4157					Y							Y																										Y	Y		
4158					Y							Y							Y							Y			Y	Y	Y					Y			Y		Y
4159					Y							Y					Y					Y							Y	Y										Y
4161					Y							Y							Y							Y			Y							Y					
4164					Y							Y	Y																							Y					
4168					Y							Y																Y								Y					
4173					Y							Y														Y		Y	Y												
4175					Y							Y																Y	Y												
4178					Y							Y																							Y	Y					
4180					Y							Y										Y												Y							
4184					Y							Y							Y																	Y					
4186					Y							Y										Y						Y		Y											
4188					Y							Y							Y																					Y	
4189					Y							Y							Y			Y				Y		Y	Y	Y						Y			Y	Y	Y	Y
4206						Y	Y					Y																												Y
4218						Y		Y			Y	Y		Y																							Y	Y			
4228						Y		Y	Y			Y																													
4240						Y		Y				Y														Y								Y							
4267						Y		Y				Y																Y													
4274						Y		Y				Y											Y				Y	Y										Y			
4276						Y		Y				Y											Y				Y	Y													
4279						Y			Y			Y																								Y					
4285						Y			Y	Y		Y															Y	Y	Y	Y					Y						
4302						Y			Y			Y																Y	Y												
4304						Y			Y			Y							Y								Y												Y			
4322						Y				Y		Y																	Y	Y										Y	Y
4325						Y				Y		Y														Y															
4506						Y						Y											Y						Y	Y											
4508						Y						Y							Y																	Y	Y		Y		
4509						Y						Y				Y														Y									Y		
4510						Y						Y							Y				Y					Y								Y					
4511						Y						Y										Y								Y									Y		Y
4512						Y						Y										Y																		Y
4513						Y						Y																Y	Y							Y				
4515						Y						Y																	Y	Y										
4522						Y						Y											Y				Y	Y								Y		Y	Y		
4526						Y						Y							Y										Y	Y						Y					

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichornutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
4529						Y						Y										Y					Y														
4538						Y						Y		Y							Y		Y				Y														
4539						Y						Y	Y									Y				Y	Y	Y	Y		Y					Y	Y			Y	Y
4540						Y						Y															Y											Y		
4541						Y						Y											Y				Y													
4544						Y						Y													Y					Y										
4546						Y						Y															Y	Y	Y											
4547						Y						Y	Y			Y																						Y		
4550						Y						Y							Y			Y														Y				
4554						Y						Y							Y									Y	Y							Y				
4555						Y						Y														Y											Y	Y		Y
4559						Y						Y															Y										Y			
4563						Y						Y														Y										Y		Y			
4567						Y						Y							Y								Y		Y							Y				Y
4568						Y						Y														Y								Y			Y			
4569						Y						Y																				Y		Y	Y				Y	Y
4570						Y						Y	Y															Y												
4572						Y						Y		Y									Y				Y										Y				
4574						Y						Y							Y			Y					Y	Y	Y							Y		Y			Y
4575						Y						Y																Y	Y											
4578						Y						Y	Y	Y																										
4579						Y						Y							Y				Y				Y	Y	Y											
4580						Y						Y		Y													Y	Y	Y							Y					
4590						Y						Y														Y	Y										Y				
4592						Y						Y		Y									Y													Y					
4597						Y						Y	Y													Y			Y	Y				Y							
4599						Y						Y															Y												Y		
4601						Y						Y															Y	Y										Y			
4602							Y	Y				Y																									Y			
4603							Y	Y				Y														Y				Y							Y			
4632							Y		Y			Y																						Y	Y			Y	Y		
4656							Y					Y			Y							Y				Y		Y	Y									Y		
4660							Y					Y	Y																										Y		
4662							Y					Y																Y	Y											
4666							Y					Y	Y						Y																					
4670							Y					Y											Y				Y		Y						Y				Y			
4671							Y					Y	Y						Y								Y	Y										Y			
4672							Y					Y																		Y									Y		
4673							Y					Y														Y			Y												
4674							Y					Y															Y											Y			
4675							Y					Y														Y			Y												
4678							Y					Y	Y																Y												
4682							Y					Y																Y											Y		
4683							Y					Y																				Y									
4684							Y					Y										Y							Y												Y
4686							Y					Y		Y		Y																									
4687							Y					Y										Y				Y				Y										Y		

	A.thaliana	*A.fumigatus*	*A.anophagefferens*	*B.natans*	*C.elegans*	*C.reinhardtii*	*C.parvum*	*C.merolae*	*D.discoidium*	*D.melanogaster*	*E.siliculosus*	*E.huxleyi*	*E.histolytica*	*E.coli*	*F.serratus*	*G.lamblia*	*G.theta*	*H.triquetra*	*H.sapiens*	*K.brevis*	*K.micrum*	*L.infantum*	*O.sativa*	*O.tauri*	*P.marina*	*P.tetraurelia*	*P.trichornutum*	*P.patens*	*P.ramorum*	*P.sojae*	*P.falciparum*	*P.haitanensis*	*P.yezoensis*	*R.filosa*	*S.cerevisiae*	*S.pombe*	*S.purpuratus*	*S.6803*	*T.pseudonana*	*T.gondii*	*T.vaginalis*	*T.brucei*
4689							Y					Y																	Y										Y		
4690							Y					Y											Y								Y											
4691							Y					Y															Y	Y	Y													
4692							Y					Y				Y							Y																Y	Y			
4697							Y					Y	Y														Y												Y				
4698							Y					Y							Y				Y			Y		Y			Y				Y							
4700							Y					Y				Y									Y											Y							
4702							Y					Y														Y			Y													
4703								Y	Y			Y															Y															
4705								Y	Y			Y				Y																									Y		
4706								Y	Y			Y																Y	Y													
4711								Y	Y			Y										Y	Y					Y	Y	Y												
4725								Y	Y			Y		Y					Y			Y	Y				Y	Y	Y	Y			Y		Y	Y	Y		Y			
4737								Y		Y		Y				Y																										
4752								Y			Y	Y																										Y	Y			
4753								Y			Y	Y										Y																	Y			Y
4754								Y			Y	Y							Y								Y		Y	Y						Y		Y			
4755								Y			Y	Y																Y										Y		Y		
4757								Y				Y											Y				Y									Y		Y				
4758								Y				Y															Y				Y		Y					Y				
4759								Y				Y											Y					Y										Y				
4761								Y				Y															Y												Y				
4765								Y				Y																													Y		
4766								Y				Y		Y					Y										Y	Y							Y						
4768								Y				Y											Y					Y														
4769								Y				Y																Y							Y			Y			Y		
4771								Y				Y														Y			Y													
4777								Y				Y							Y								Y										Y			Y	Y		
4780								Y				Y																										Y	Y			
4782								Y				Y	Y						Y									Y	Y												Y	
4786								Y				Y														Y		Y														
4791								Y				Y	Y																Y	Y												
4792								Y				Y															Y												Y				
4793								Y				Y											Y					Y														
4795								Y				Y														Y	Y												Y				
4796								Y				Y																						Y					Y			
4798								Y				Y															Y												Y				
4802								Y				Y											Y					Y											Y				
4803								Y				Y															Y												Y				
4806								Y				Y																								Y						Y
4807								Y				Y															Y							Y								
4808								Y				Y							Y										Y	Y												Y
4811								Y				Y															Y						Y									
4812								Y				Y															Y												Y				
4816								Y				Y				Y			Y							Y											Y	Y					
4817								Y				Y	Y			Y						Y								Y	Y										Y		
4819								Y				Y															Y	Y											Y				


A.thaliana
A.fumigatus
A.anophagefferens
B.natans
C.elegans
C.reinhardtii
C.parvum
C.merolae
D.discoidium
D.melanogaster
E.siliculosus
E.huxleyi
E.histolytica
E.coli
F.serratus
G.lamblia
G.theta
H.triquetra
H.sapiens
K.brevis
K.micrum
L.infantum
O.sativa
O.tauri
P.marina
P.tetraurelia
P.trichomutum
P.patens
P.ramorum
P.sojae
P.falciparum
P.haitanensis
P.yezoensis
R.filosa
S.cerevisiae
S.pombe
S.purpuratus
S.6803
T.pseudonana
T.gondii
T.vaginalis
T.brucei
4821
4824
4825
4829
4832
4836
4837
4838
4840
4842
4844
4845
4851
4852
4856
4859
4871
4880
4881
4883
4886
4888
4891
4892
4893
4895
4897
4898
4902
4904
4905
4906
4908
4909
4910
4913
4917
4920
4922
4924
4925
4926
4928
4929
4930
4933
4934


A.thaliana
A.fumigatus
A.anophagefferens
B.natans
C.elegans
C.reinhardtii
C.parvum
C.merolae
D.discoidium
D.melanogaster
E.siliculosus
E.huxleyi
E.histolytica
E.coli
F.serratus
G.lamblia
G.theta
H.triquetra
H.sapiens
K.brevis
K.micrum
L.infantum
O.sativa
O.tauri
P.marina
P.tetraurelia
P.trichomutum
P.patens
P.ramorum
P.sojae
P.falciparum
P.haitanensis
P.yezoensis
R.filosa
S.cerevisiae
S.pombe
S.purpuratus
S.6803
T.pseudonana
T.gondii
T.vaginalis
T.brucei
4938
4939
4940
4941
4942
4944
4946
4947
4949
4950
4951
4954
4957
4959
4961
4962
4964
4965
4966
4967
4969
4971
4972
4973
4974
4977
4978
4979
4983
4987
4988
4989
4991
4992
4993
4994
4996
4998
5001
5010
5011
5012
5013
5014
5015
5017
5018


A.thaliana
A.fumigatus
A.anophagefferens
B.natans
C.elegans
C.reinhardtii
C.parvum
C.merolae
D.discoidium
D.melanogaster
E.siliculosus
E.huxleyi
E.histolytica
E.coli
F.serratus
G.lamblia
G.theta
H.triquetra
H.sapiens
K.brevis
K.micrum
L.infantum
O.sativa
O.tauri
P.marina
P.tetraurelia
P.trichomutum
P.patens
P.ramorum
P.sojae
P.falciparum
P.haitanensis
P.yezoensis
R.filosa
S.cerevisiae
S.pombe
S.purpuratus
S.6803
T.pseudonana
T.gondii
T.vaginalis
T.brucei
5019
5020
5021
5022
5025
5026
5027
5029
5030
5033
5034
5035
5036
5037
5040
5042
5043
5045
5046
5049
5050
5051
5052
5054
5055
5060
5062
5064
5066
5069
5070
5071
5072
5076
5083
5085
5087
5091
5093
5094
5097
5098
5100
5101
5102
5105
5106
5108

References

Abrahamsen MS, Templeton TJ, Enomoto S, Abrahante JE, Zhu G, Lancto CA, Deng M, Liu C, Widmer G, Tzipori S, Buck GA, Xu P, Bankier AT, Dear PH, Konfortov BA, Spriggs HF, Iyer L, Anantharaman V, Aravind L, Kapur V (2004) Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science*. 304: 441-445.

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers YH, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Gabor GL, Abril JF, Agbayani A, An HJ, Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadieu E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, de Pablos B, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Heiman TJ, Hernandez JR, Houck J, Hostin D, Houston KA, Howland TJ, Wei MH, Ibegwam C, Jalali M,

Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, McIntosh TC, McLeod MP, McPherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nusskern DR, Pacleb JM, Palazzolo M, Pittman GS, Pan S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Saunders RD, Scheeler F, Shen H, Shue BC, Sidén-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas R, Tector C, Turner R, Venter E, Wang AH, Wang X, Wang ZY, Wassarman DA, Weinstock GM, Weissenbach J, Williams SM, WoodageT, Worley KC, Wu D, Yang S, Yao QA, Ye J, Yeh RF, Zaveri JS, Zhan M, Zhang G, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith HO, Gibbs RA, Myers EW, Rubin GM, Venter JC (2000) The genome sequence of *Drosophila melanogaster*. *Science*. 287:2185-95.

Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, and Taylor MF (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol*. 52: 399-451.

Alexander E, Stock A, Breiner H-W, Behnke A, Bunge J, Yakimov MM, and Stoeck T (2009) Microbial eukaryotes in the hypersaline anoxic L'Atalante deep-sea basin. *Env Microbiol.* 11: 360-381.

Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25: 3389-3402.

Anderson, RA (2004) Biology and systematics of heterokont and haptophyte algae. *Am J Bot.* 91: 1508-1522.

Appleton PL, Vickerman K (1998) *In vitro* cultivation and developmental cycle in culture of a parasitic dinoflagellate (*Hematodinium* sp.) associated with mortality of the Norway lobster (*Nephrops norvegicus*) in British waters. *Parasitology.* 116: 115-130.

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature. 408:796-815.

Armbrust, EV, Berges, JA, Bowler, C, Green, BR, Martinez, D, Putnam, NH, Zhou, SG, Allen, AE, Apt, KE, Bechner, M, Brzezinski, MA, Chaal, BK, Chiovitti, A, Davis, AK, Demarest, MS, Detter, JC, Glavina, T, Goodstein, D, Hadi, MZ, Hellsten, U, Hildebrand, M, Jenkins, BD, Jurka, J, Kapitonov, VV, Kroger, N, Lau, WWY, Lane, TW, Larimer, FW, Lippmeier, JC, Lucas, S, Medina, M, Montsant, A, Obornik, M, Parker, MS, Palenik, B, Pazour, GJ, Richardson, PM, Rynearson,

TA, Saito, MA, Schwartz, DC, Thamatrakoln, K, Valentin, K, Vardi, A, Wilkerson, FP, and Rokhsar, DS The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. *Science*. 306: 79-86.

Aury JM, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, Ségurens B, Daubin V, Anthouard V, Aiach N, Arnaiz O, Billaut A, Beisson J, Blanc I, Bouhouche K, Câmara F, Duharcourt S, Guigo R, Gogendeau D, Katinka M, Keller AM, Kissmehl R, Klotz C, Koll F, Le Mouël A, Lepère G, Malinsky S, Nowacki M, Nowak JK, Plattner H, Poulain J, Ruiz F, Serrano V, Zagulski M, Dessen P, Bétermier M, Weissenbach J, Scarpelli C, Schächter V, Sperling L, Meyer E, Cohen J, Wincker P. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. (2006) *Nature*. 444(7116): 171-8.

Bachvaroff TR, Sanchez Puerta MV, Delwiche CF. (2005) Chlorophyll c-containing plastid relationships based on analyses of a multigene data set with all four chromalveolate lineages. *Mol Biol Evol*. 22: 1772-82.

Bachvaroff TR, Sanchez-Puerta MV, and Delwiche CF (2006) Rate variation as a function of gene origin in plastid-derived genes of peridinin-containing dinoflagellates. *J Mol Evol*. 62: 42-52.

Baum J, Gilberger T-W, Frischknecht F, and Meissner M (2008) Host-cell invasion by malaria parasites: insights from *Plasmodium* and *Toxoplasma.* *Trends Parasitol*. 24: 557-563.

Bhattacharya, D, Medlin, L, Wainright, PO, Ariztia, EV, Bibeau, C, Stickel, SK, and Sogin, ML (1992) Algae containing chlorophylls a + c are paraphyletic: molecular evolutionary analysis of the chromophyta. *Evolution* 46: 1801-1817.

Bhattacharya, D, Stickel, SK, and Sogin, ML (1993) Isolation and molecular phylogenetic analysis of actin-coding regions from *Emiliania huxleyi*, a prymnesiophyte alga, by reverse transcriptase and PCR methods. *Mol. Biol. Evol*. 10: 689-703.

Bhaud Y, Guillebault D, Lennon J, Defacque H, Soyer-Gobillard MO, Moreau H (2000) Morphology and behaviour of dinoflagellate chromosomes during the cell cycle and mitosis. *J Cell Sci.* 113: 1231-1239.

Bjornland T, Haxo FT, and Liaaen-Jensen S (2003) Carotenoids of the Florida red tide dinoflagellate *Karenia brevis*. *Biochem Syst Ecol*. 31: 1147-1162.

Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*. 277: 1453-62

Bodyl, A (2005) Do plastid-related characters support the chromalveolate hypothesis. *J Phycol*. 41: 712-719.

Bouck, GB (1969) Extracellular microtubules. The origin, structure, and attachment of flagellar hairs in *Fucus* and *Ascophyllum* antherozoids. *J Cell Biol*. 40: 446-460.

Bouck GB, Sweeney BM (1966) The fine structure and ontogeny of trichocysts in marine dinoflagellates. *Protoplasma.* 61: 205-223.

Bourrelly, P (1965) La classification des Chrysophycées: ses problèmes. *Revue algol*. 1: 56-60.

Bowler, C, Allen, AE, Badger, JH, Grimwood, J, Jabbari, K, Kuo, A, Maheswari, U, Martens, C, Maumus, F, Otillar, RP, Rayko, E, Salamov, A, Vandepoele, K, Beszteri, B, Gruber, A, Heijde, M, Katinka, M, Mock, T, Valentin, K, Verret, F, Berges, JA, Brownlee, C, Cadoret, JP, Chiovitti, A, Choi, CJ, Coesel, S, De Martino, A, Detter, JC, Durkin, C, Falciatore, A, Fournet, J, Haruta, M, Huysman, MJJ, Jenkins, BD, Jiroutova, K, Jorgensen, RE, Joubert, Y, Kaplan, A, Kroger, N, Kroth, PG, La Roche, J, Lindquist, E, Lommer, M, Martin-Jezequel, V, Lopez, PJ, Lucas, S, Mangogna, M, McGinnis, K, Medlin, LK, Montsant, A, Oudot-Le Secq, MP, Napoli, C, Obornik, M, Parker, MS, Petit, JL, Porcel, BM, Poulsen, N, Robison, M, Rychlewski, L, Rynearson, TA, Schmutz, J, Shapiro, H, Siaut, M, Stanley, M, Sussman, MR, Taylor, AR, Vardi, A, von Dassow, P, Vyverman, W, Willis, A, Wyrwicz, LS, Rokhsar, DS, Weissenbach, J, Armbrust, EV, Green, BR, Van De Peer, Y, and Grigoriev, IV (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom Genomes. *Nature*. 456: 239-244.

Brugerolle G (2002) *Cryptophagus subtilis*: a new parasite of cryptophytes affiliated with the Perkinsozoa lineage. *Eur J Protistol.* 37: 379-390.

Brugerolle G (2003) Apicomplexan parasite *Cryptophagus* renamed *Rastrimonas* gen. nov.. *Eur J Protistol.* 39: 101-101.

Brugerolle, G (2002) *Colpodella vorax*: ultrastructure, predation, life-cycle mitosis, and phylogenetic relationships. *Eur J Protistol.* 38: 113-125.

Burki F, Nikolaev SI, Bolivar I, Guiard J, Pawlowski J (2006) Analysis of expressed sequence tags from naked foraminiferan Reticulomyxa filosa. Genome. 49: 882-7.

Burki, F, Shalchian-Tabrizi, K, Minge, M, Skjaeveland, A, Nikolaev, SI, Jakobsen, KS, and Pawlowski, J (2007) Phylogenomics reshuffles the eukaryotic supergroups. *PloS ONE* 2: e790. doi:10.1371/journal.pone.0000790.

Burki, F, Shalchian-Tabrizi K, Pawlowski, J (2008) Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. *Biol Lett.* 4: 366-369.

Cachon J (1964) Contribution a l□fétude des Péridiniens parasites cytology, cycles évolutifs. *Ann Sci Nat Zool.* 12: 1-158.

Cachon J, Cachon M (1969) Ultrastructures des Amoebophryidae (Péridiniens Duboscquodinida) I. Manifestations des rapports entre l'hote et le parasite. *Protistologica* 5: 535-547.

Cachon J, Cachon M (1970) Ultrastructures des Amoebophryidae (Péridiniens Duboscquodinida) II. Systèms Atractrophoriens et microtubulaires; leur interventions dans la mitose. *Protistologica*. 6: 57-70.

Cachon J and Cachon M (1987) Parasitic Dinoflagellates. *The Biology of Dinoflagellates*. Blackwell Scientific, Boston, pp 571-610.

Calado AJ, Craveiro SC, and Moestrup Æ (1998) Taxonomy and utrastructure of a freshwater, heterotrophic *Amphidinium* (Dinophyceae) that feeds on unicellular protists. *J Phycol.* 34: 536-554.

Carruthers VB and Sibley LD (1997) Sequential protein secretion from three distinct organelles of *Toxoplasma gondii* accompanies invasion of human fibroblasts. *Eur J Cell Biol.* 73: 114-123.

Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark UC, Besteiro S, Sicheritz-Ponten T, Noel CJ, Dacks JB, Foster PG, Simillion C, Van de Peer Y, Miranda-Saavedra D, Barton GJ, Westrop GD, Müller S, Dessi D, Fiori PL, Ren Q, Paulsen I, Zhang H, Bastida-Corcuera FD, Simoes-Barbosa A, Brown MT, Hayes RD, Mukherjee M, Okumura CY, Schneider R, Smith AJ, Vanacova S, Villalvazo M, Haas BJ, Pertea M, Feldblyum TV, Utterback TR, Shu CL, Osoegawa K, de Jong PJ, Hrdy I, Horvathova L, Zubacova Z, Dolezal P, Malik SB, Logsdon JM Jr, Henze K, Gupta A, Wang CC, Dunne RL, Upcroft JA, Upcroft P, White O, Salzberg SL, Tang P, Chiu CH, Lee YS, Embley TM, Coombs GH, Mottram JC, Tachezy J, Fraser-

Liggett CM, Johnson PJ (2007) Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis. Science.* 315: 207-212.

Castillo-Davis CI and Hartl DL (2003) GeneMerge--post-genomic analysis, data mining, and hypothesis testing. *Bioinformatics*. 19: 891-892.

Cavalier-Smith T and Chao EE (2004) Protalveolate phylogeny and systematics and the origins of Sporozoa and dinoflagellates (phylum Myzozoa nom. nov.). *Eur J Protistol.* 40: 185-212.

Cavalier-Smith T (1981) Eukaryote Kingdoms: seven or nine? *BioSystems*. 14: 461-481.

Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryotic family tree. *J Eukaryot Microbiol*. 46: 347-366.

Cavalier-Smith T (2006) Cell evolution and earth history: stasis and revolution. *Phil Trans R Soc B.* 361: 969-1006.

C. elegans Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*. 282: 2012-8.

Charlson RJ, Lovelock JE, Andreae MO, Warren SG (1987) Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature.* 326: 655–661.

Cheung MK, Chu KH, Li CP, Kwan, HS, and Wong CK (2008) Genetic diversity of picoeukaryotes in a semi-enclosed harbor in the subtropical western Pacific Ocean. *Aquat Microb Ecol*. 53: 295-305.

Christensen T (1962) Alger. In TW Böcher, M Lange, T Sørensen (Eds), Botanik, 2(Systematik Botanik), 2: 1-178. Copenhagen: Munksgaard.

Clay BL and Kugrens P (1999) Characterization of *Hemiselmis amylosa* sp. nov. and phylogenetic placement of the blue-green cryptomonads *H. amylosa* and *Falcomonas daucoides*. *Protist*. 150: 297-310.

Coats DW (1988) *Duboscquella cachoni* n. sp., a parasitic dinoflagellate lethal to its tintinnine host *Eutintinnus pectinis*. *J Protozool.* 35: 607-617.

Coats DW (1999) Parasitic lifestyles of marine dinoflagellates. *J Eukaryot Microbiol*. 46: 402-409.

Coats DW, Adam EJ, Gallegos CL, Hedrick S (1996) Parasitism of photosynthetic dinoflagellates in a shallow subestuary of Chesapeake Bay, USA. *Aquat Microb Ecol.* 11: 1-9.

Coats DW, Bockstahler KR (1994) Occurrence of the parasitic dinoflagellate *Amoebophrya ceratii* in Chesapeake Bay populations of *Gymnodinium sanguineum*. *J Eukaryot Microbiol*. 41: 586-593.

Coats DW, Park MG (2002) Parasitism of photosynthetic dinoflagellates by three strains of *Amoebophrya* (Dinophyta): Parasite survival, infectivity, generation time, and host specificity. *J Phycol* 38: 520-528.

Daubin V, Gouy M, Perrière G (2002) A phylogenomic approach to bacterial phylogeny: evidence of a core of genes sharing a common history. *Genome Research*. 12: 1080-1090.

Day JD, Hall CAS, Kemp WM, Yanez_Arancibia A (1989) *Estuarine Ecology*. John Wiley and Sons, New York, p 385.

Delwiche, CF (1999) Tracing the thread of plastid diversity through the tapestry of life. *Am Nat*. 154: S164-S177.

Delwich CF and Palmer JD (1996) Rampant horizontal transfer and duplication of Rubisco genes in eubacteria and plastids. *Mol Biol Evol*. 13: 873-882.

Deschamps P and Moreira D (2009) Signal conflict in the phylogeny of the primary photosynthetic Eukaryotes. *Mol Biol Evol*. 26: 2745-2753.

Dittami SM, Scornet D, Petit JL, Ségurens B, Da Silva C, Corre E, Dondrup M, Glatting KH, König R, Sterck L, Rouzé P, Van de Peer Y, Cock JM, Boyen C, Tonon T (2009) Global expression analysis of the brown alga *Ectocarpus siliculosus* (Phaeophyceae) reveals large-scale reprogramming of the transcriptome in response to abiotic stress. *Genome Biol*. 2009;10:R66 doi:10.1186/gb-2009-10-6-r66.

Dobrowolski JM, Sibley LD (1996) *Toxoplasma* invasion of mammalian cells is powered by the actin cytoskeleton of the parasite. *Cell*. 84: 933-939.

Dodge JD, Crawford RM (1971) Fine structure of the dinoflagellate *Oxyrrhis marina* II. the flagellar system. *Protisologica.* 7: 399-409.

Dodge JD, Greuet C (1987) Dinoflagellate ultrastructure and complex organelles. *The Biology of Dinoflagellates*. Blackwell Scientific, Boston, pp 92-142.

Douglas, SE and Penny, SL (1999) The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. *J Mol Evol*. 48: 236-244.

Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu X, Reith M, Cavalier-Smith T, Maier UG (2001) The highly reduced genome of an enslaved algal nucleus. *Nature*. 410: 1091-6.

Dubremetz JF, Garcia-Reguet N, Conseil V, Fourmaux MN (1998) Apical organelles and host-cell invasion by Apicomplexa. *Int J Parasitol*. 28: 1007-1013.

Edgar, RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research. 32: 1792-1797.

Edwards, SV (2009) Is a new and general theory of molecular systematics emerging? *Evolution*. 63: 1-19.

Eichinger L, Pachebat JA, Glöckner G, Rajandream MA, Sucgang R, Berriman M, Song J, Olsen R, Szafranski K, Xu Q, Tunggal B, Kummerfeld S, Madera M, Konfortov BA, Rivero F, Bankier AT, Lehmann R, Hamlin N, Davies R, Gaudet P, Fey P, Pilcher K, Chen G, Saunders D, Sodergren E, Davis P, Kerhornou A, Nie X, Hall N, Anjard C, Hemphill L, Bason N, Farbrother P, Desany B, Just E, Morio T, Rost R, Churcher C, Cooper J, Haydock S, van Driessche N, Cronin A, Goodhead I, Muzny D, Mourier T, Pain A, Lu M, Harper D, Lindsay R, Hauser H, James K, Quiles M, Madan Babu M, Saito T, Buchrieser C, Wardroper A, Felder M, Thangavelu M, Johnson D, Knights A, Loulseged H, Mungall K, Oliver K, Price C, Quail MA, Urushihara H, Hernandez J, Rabbinowitsch E, Steffen D, Sanders M, Ma J, Kohara Y, Sharp S, Simmonds M, Spiegler S, Tivey A, Sugano S, White B, Walker D, Woodward J, Winckler T, Tanaka Y, Shaulsky G, Schleicher M, Weinstock G, Rosenthal A, Cox EC, Chisholm RL, Gibbs R, Loomis WF, Platzer M, Kay RR, Williams J, Dear PH, Noegel AA, Barrell B, Kuspa A (2005) The genome of the social amoeba *Dictyostelium discoideum*. *Nature*. 435: 43-57.

Eisen JA, Coyne RS, Wu M, Wu D, Thiagarajan M, Wortman JR, Badger JH, Ren Q, Amedeo P, Jones KM, Tallon LJ, Delcher AL, Salzberg SL, Silva JC, Haas BJ, Majoros WH, Farzad M, Carlton JM, Smith RK Jr, Garg J, Pearlman RE, Karrer KM, Sun L, Manning G, Elde NC, Turkewitz AP, Asai DJ, Wilkes DE, Wang Y, Cai H, Collins K, Stewart BA, Lee SR, Wilamowska K, Weinberg Z, Ruzzo WL, Wloga D, Gaertig J, Frankel J, Tsao CC, Gorovsky MA, Keeling PJ, Waller RF, Patron NJ, Cherry JM, Stover NA, Krieger CJ, del Toro C, Ryder HF,

Williamson SC, Barbeau RA, Hamilton EP, Orias E (2006) Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biol*. 2006 4: e286.

Elliot DA, McIntosh MT, Hosgood III HD, Chen S, Zhang G, Baevova P, Joiner KA (2008) Four distinct pathways of hemoglobin uptake in the malaria parasite *Plasmodium falciparum*. *Proc Natl Acad Sci*. 105: 2463-2468.

El-Sayed NM, Myler PJ, Blandin G, Berriman M, Crabtree J, Aggarwal G, Caler E, Renauld H, Worthey EA, Hertz-Fowler C, Ghedin E, Peacock C, Bartholomeu DC, Haas BJ, Tran AN, Wortman JR, Alsmark UC, Angiuoli S, Anupama A, Badger J, Bringaud F, Cadag E, Carlton JM, Cerqueira GC, Creasy T, Delcher AL, Djikeng A, Embley TM, Hauser C, Ivens AC, Kummerfeld SK, Pereira-Leal JB, Nilsson D, Peterson J, Salzberg SL, Shallom J, Silva JC, Sundaram J, Westenberger S, White O, Melville SE, Donelson JE, Andersson B, Stuart KD, Hall N (2005) Comparative genomics of trypanosomatid parasitic protozoa. *Science*. 309: 404-9.

Erdner DL, Dyble J, Parsons ML, Stevens RC, Hubbard KA, Wrabel ML, Moore SK, Lefebvre KA, Anderson DM, Bienfang P, Bidigare RR, Parker MS, Moeller P, Brand LE, Trainer VL. (2008) Centers for Oceans and Human Health: a unified approach to the challenge of harmful algal blooms. *Environ Health*. Suppl 2:S2.

Falkowski, PG, Barber, RT, and Smetacek, V (1998) Biogeochemical controls and feedbacks on ocean primary production. *Science*. 281: 200-206.

Falkowski, PG, Katz, ME, Knoll, AH, Quigg, A, Raven, JA, Scholfield, O, and Taylor, FJR (2004) The evolution of modern eukaryotic phytoplankton. *Science*. 305: 354-360.

Fast NM, Xue L, Bingham S Keeling PJ (2002) Re-examining alveolate evolution using multiple protein molecular phylogenies. *J Eukaryot Microbiol.* 49: 30-7.

Felsenstein, J. 1989. PHYLIP -- Phylogeny Inference Package (Version 3.2). Cladistics 5: 164-166.

Frensel J and Probert I (2005) The ultrastructure and life cycle of the coastal coccolithophorid *Ochrosphaera neapolitana* (Prymnesiophyceae). *Eur J Phycol*. 40: 105-122.

Fritz L, Nass M (1992) Development of the endoparasitic dinoflagellate *Amoebophyra ceratii* within host dinoflagellate species. *J Phycol*. 28: 312-320.

Fritz-Laylin, LK, Prochnik, SE, Ginger, ML, Dacks, JB, Carpenter, ML, Field, MC, Kuo, A, Paredez, A, Chapman, J, Pham, J, Shu, SQ, Neupane, R Cipriano, M, Mancuso, J, Tu, H, Salamov, A, Lindquist, E, Shapiro, H, Lucas, S, Grigoriev, IV, Cande, WZ, Fulton, C, Rokhsar, DS, and Dawson, SC (2010) The Genome of *Naegleria gruberi* Illuminates Early Eukaryotic Versatility. Cell. 140: 631-642.

Gajadhar AA, Marquardt WC, Hall R, Gunderson J, Ariztia-Carmona EV, Sogin ML (1991) Ribosomal RNA sequences of *Sarcocystis muris*, *Theileria annulata*

and *Crypthecodinium cohnii* reveal evolutionary relationships among apicomplexans, dinoflagellates, and ciliates. *Mol Biochem Parasitol.* 45: 147-54.

Gajria B, Bahl A, Brestelli J, Dommer J, Fischer S, Gao X, Heiges M, Iodice J, Kissinger JC, Mackey AJ, Pinney DF, Roos DS, Stoeckert CJ Jr, Wang H, Brunk BP (2008) ToxoDB: an integrated *Toxoplasma gondii* database resource. *Nucleic Acids Res*. 36:D553-6.

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419: 498-511.

Gantt E (1975) Phycobilisomes: light-harvesting pigment complexes. *Bioscience*. 25: 781-788.

Gantt E, Edwards MR, and Provasoli L (1971) Chloroplast structure of the Cryptophyceae. evidence for phycobiliproteins with intrathylakoid spaces. *J Cell Biol*. 48: 280-290.

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K,

Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, Barrell B (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature.* 419: 498-511.

Gibbs, SP (1981) The Chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae. *Ann NY Acad Sci*. 361: 193-208.

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. *Science*. 274: 546, 563-7.

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science*. 296: 92-100.

Gould SB, Sommer MS, Hadfi K, Zauner S, Kroth PG, and Maier U-G (2006) Protein targeting into the complex plastid of cryptophytes. *J Mol Biol*. 62: 674-681.

Graham, LE and Wilcox, LW (2000) *Algae*. Prentice Hall. NJ, USA, 180-197, 233-268.

Green, JC, Course, PA, and Tarran, GA (1996) The life-cycle of *Emiliania huxleyi*: a brief outline and a study of relative ploidy levels analysed by flow cytometry. *J Marine Syst*. 9: 33-44.

Grey MW, Burger G, and Lang, BF (1999) Mitochondrial Evolution. *Science*. 283: 1476-1481.

Groisillier A, Massana R, Valentin K, Vaulot D, and Guillou, L. (2006). Genetic diversity and habitats of two enigmatic marine alveolate lineages. *Aquat Microb Ecol*. 42(3): 277-291.

Gruber A, Vugrinec S, Hempel F, Gould SB, Maier, U-G, and Kroth (2007) Protein targeting into complex diatom plastids: functional characterization of a specific targeting motif. *Plant Mol Biol*. 64: 519-530.

Hackett JD, Scheetz TE, Yoon HS, Soares MB, Bonaldo MF, Casavant TL, Bhattacharya D (2005) Insights into a dinoflagellate genome through expressed sequence tag analysis. *BMC Genomics*. 6: 80. doi:10.1186/1471-2164-6-80.

Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rümmele SE, and Bhattacharya D (2007) Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. *Mol. Biol. Evol*. 24: 1702–1713.

Haidak DJ, Mathews CK, and Sweeney BM (1966) Pigment protein complex from *Gonyaulax*. *Science*. 152: 212-213.

Hakansson S, Charron AJ, and Sibley LD. (2001). *Toxoplasma* evacuoles: a two-step process of secretion and fusion forms the parasitophorous vacuole. *EMBO Journal*. 20: 3132-3144.

Haldar K, Kamoun S, Hiller NL, Bhatacharje S, and van Ooij C (2006) Common infection strategies of pathogenic eukaryotes. *Nature*. 4: 922-931.

Hall N, Karras M, Raine JD, Carlton JM, Kooij TW, Berriman M, Florens L, Janssen CS, Pain A, Christophides GK, James K, Rutherford K, Harris B, Harris D, Churcher C, Quail MA, Ormond D, Doggett J, Trueman HE, Mendoza J, Bidwell SL, Rajandream MA, Carucci DJ, Yates JR 3rd, Kafatos FC, Janse CJ, Barrell B, Turner CM, Waters AP, Sinden RE (2005) A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science* 307: 82-86.

Harada, A, Ohtsuka, S, Horiguchi, T (2007) Species of the parasitic genus *Duboscquella* are members of the enigmatic Marine Alveolate Group I. *Protist.* 158: 337-347.

Harb OS, Chatterjee B, Fraunhoz MJ, Crawford MJ, Nishi M, and Roos DS (2004) Multiple functionally redundant signals mediate targeting to the apicoplast in the apicomplexan parasite *Toxoplasma gondii*. *Eukaryotic Cell*. 3: 663-674.

Hardham AR (2007) Cell biology of plant-oomycete interactions. *Cell Microbiol*. 9: 31-39.

Hausmann, K, Hülsmann, N (1996) *Protozoology*, 2nd edition Thieme Medical Publishers, New York pp 87-107.

Hibberd, DJ (1976) The ultrastructure and taxonomy of the Chrysophyceae and Prymnesiophyceae (Haptophyceae): a survey with some new observations on the ultrastructure of the Chrysophyceae. *Botanical Journal of the Linnean Society*. 72: 55-80.

Hohfeld, I, Melkonian, M (1995) Ultrastructure of the flagellar apparatus of *Noctiluca miliaris* Suriray swarmers (Dinophyceae). *Phycologia.* 34: 508-513.

Hornik (2009) The R FAQ http://www.r-project.org/ ISBN 3-900051-08-9.

Huang, X. and Madan, A. (1999) CAP3: A DNA sequence assembly program. *Genome Res.*, 9, 868-877.

Iglesias-Prieto R, Govind NS, Trench RK (1993) Isolation and characterization of three membrane-bound chlorophyll-protein complexes from four dinoflagellate species. *Philos T R Soc B*. 340: 381-392.

Iglesias-Rodriguez, MD, Halloran, PR, Rosalind E. M. Rickaby, REM, Hall, IR, Colmenero-Hidalgo, E, Gittins, JR, Green, DRH, Tyrrell, T, Gibbs, SJ, von Dassow, P, Rehm, E, Armbrust, EV, Boessenkool, KP (2008) Phytoplankton calcification in a high-CO_2 world. *Science*. 320: 336-340.

International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature*. 409: 860-921.

Jain R, Rivera MC, and Lake, JA (1999) Horizontal gene transfer among genomes: the complexity hypothesis. *Proc Natl Acad Sci*. 96: 3801-3806.

Jeffroy, O, Brinkmann, H, Delsuc, F, and Philippe, H (2006) Phylogenomics: the beginning of incongruence. *Trends Genet*. 22: 225-231.

Jékely G (2009) Evolution of phototaxis. *Phil. Trans. R. Soc. B*. 364: 2795-2808.

Jones DT, Taylor WR, and Thornton JM (1992) A new approach to protein fold recognition. *Nature*. 358: 86-89.

Kawachi M, Inouye I, Maeda O, Chihara M (1991) The haptonema as a food-capturing device: observations on *Chrysochromulina hirta* (Prymnesiophyceae). *Phycologia*. 30: 563-573.

Keeling, PJ and Palmer, JD (2008) Horizontal gene transfer in eukaryotic evolution. *Nature Reviews Genetics*. 9: 605-618.

Keeling, PJ, Burger, G, Durnford, D.G. Lang, BF, Lee, RW, Pearlman, RE, Roger, AJ, and Gray, MW (2005) The Tree of Eukaryotes. *Trends Ecol Evol*. 20: 670-676.

Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S, Kimura T, Hosouchi T, Matsuno A, Muraki A, Nakazaki N, Naruo K, Okumura S, Shimpo S, Takeuchi C, Wada T, Watanabe A, Yamada M, Yasuda M, Tabata S (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* 3: 109-36.

Kim S (2006) Patterns in host range for two strains of *Amoebophrya* (Dinophyta) infecting thecate dinoflagellates: *Amoebophrya* spp. Ex *Alexandrium affine* and ex *Gonyaulax polygramma*. *J Phycol*. 42: 1170-1173.

Kim, S, Park, MG, Yih, W, Coats, DW (2004) Infection of the bloom-forming thecate dinoflagellates *Alexandrium affine* and *Gonyaulax spinifera* by two strains of *Amoebophrya* (Dinophyta). *J Phycol*. 40: 815-822.

Klaveness D (1972) *Coccolithus huxleyi* (Lohmann) Kamptner I. – Morphological investigations on the vegetative cell and the process of coccolith formation. *Protistologica*. 8: 335-346.

Kühn, SF and Medlin, LK (2005) The systematic position of the parasitoid marine dinoflagellate *Paulsenella vonstoschii* (Dinophyceae) inferred from nuclear-encoded small subunit ribosomal DNA. *Protist.* 156: 393-398.

Kuvardina, ON, Leander, BS, Aleshin, VV, Myl'nikov, AP, Keeling, PJ, Simdyanov, TG (2002) The phylogeny of colpodellids (alveolata) using small subunit rRNA gene sequences suggests they are the free-living sister group to apicomplexans. *J Eukaryot Microbiol* 49: 498-504.

Lang M, Apt KE, and Kroth PG (1998) Protein transport into “complex” diatom plastids utilizes two different targeting signals. *J Biol Chem*. 273: 30973-30978.

Le Corguillé G, Pearson G, Valente M, Viegas C, Gschloessl B, Corre E, Bailly X, Peters AF, Jubin C, Vacherie B, Cock JM, Leblanc C. (2009) Plastid genomes of two brown algae, *Ectocarpus siliculosus* and *Fucus vesiculosus*: further insights on the evolution of red-algal derived plastids. *BMC Evol Biol*. 9:253 doi:10.1186/1471-2148-9-253.

Leadbeater, B, and Dodge, JD (1967) An Electron Microscope Study of Dinoflagellate Flagella. *J Gen Microbiol.* 46: 305-314.

Leadbeater, BS (1971) Intracellular origin of flagellar hairs in dinoflagellate *Woloszynskia micra* Leadbeater & Dodge. *J Cell Sci*. 9: 443-451.

Leander, BS Hoppenrath, M (2008) Ultrastructure of a novel tube-forming, intracellular parasite of dinoflagellates: *Parvilucifera prorocentri* sp nov (Alveolata, Myzozoa). *Eur J Protistol.* 44: 55-70.

Leander, BS Keeling, PJ (2004) Early evolutionary history of dinoflagellates and apicomplexans (Alveolata) as inferred from hsp90 and actin phylogenies. J Phycol 40: 341-350.

Leander, BS Keeling, PJ (2003) Morphostasis in alveolate evolution. *Trends Ecol Evol.* 18: 395-402.

Leander, BS, Kuvardina, ON, Aleshin, VV, Mylnikov, AP, Keeling, PJ (2003) Molecular phylogeny and surface morphology of *Colpodella edax* (Alveolata): Insights into the phagotrophic ancestry of apicomplexans. *J Eukaryot Microbiol.* 50: 334-340.

Learngaramkul P, Petmitr S, Krungkrai SR, Prapunwattana, P, and Krungkrai, J (1999) Molecular characterization of mitochondria in asexual and sexual blood stages of *Plasmodium falciparum*. *Mol Cell Biol Res Commun.* 2: 15-20.

Lee RE (1989) *Phycology: second edition*. Cambridge University Press. UK.

Leedale, GF, Leadbeater, BS, Massalski, A (1970) Intracellular origin of flagellar hairs in Chrysophyceae and Xanthophyceae. *J Cell Sci*. 6: 701-719.

Leigh, JW, Susko, E, Baumgartner, M, and Roger, AJ (2008) Testing congruence in phylogenomic analysis. Syst. Biol. 57: 104-115.

Lewin RA (1992) What the haptonema is for. *Nature*. 356: 195-196.

Lidie KB, Ryan JC, Barbier M, Van Dolah FM. (2005) Gene expression in Florida red tide dinoflagellate *Karenia brevis*: analysis of an expressed sequence tag library and development of DNA microarray. *Mar Biotechnol (NY).* 7: 481-93

Lim L and McFadden I (2010) The evolution, metabolism and functions of the apicoplast. *Philos Trans R Soc Lond B.* 365: 749-763.

Litaker, RW, Tester, PA, Colorni, A, Levy, MG, and Noga, EJ (1999) The phylogenetic relationship of *Pfiesteria piscicida*, *Cryptoperidinopsoid* sp. *Amyloodinoum ocellatum* and a *Pfiesteria*-like dinoflagellate to other dinoflagellates and apicomplexans. *J Phycol*. 35: 1379-1389.

Loftus B, Anderson I, Davies R, Alsmark UC, Samuelson J, Amedeo P, Roncaglia P, Berriman M, Hirt RP, Mann BJ, Nozaki T, Suh B, Pop M, Duchene M, Ackers J, Tannich E, Leippe M, Hofer M, Bruchhaus I, Willhoeft U, Bhattacharya A, Chillingworth T, Churcher C, Hance Z, Harris B, Harris D, Jagels K, Moule S, Mungall K, Ormond D, Squares R, Whitehead S, Quail MA, Rabbinowitsch E, Norbertczak H, Price C, Wang Z, Guillén N, Gilchrist C, Stroup SE, Bhattacharya S, Lohia A, Foster PG, Sicheritz-Ponten T, Weber C, Singh U, Mukherjee C, El-Sayed NM, Petri WA Jr, Clark CG, Embley TM, Barrell B, Fraser CM, Hall N (2005) The genome of the protist parasite *Entamoeba histolytica*. *Nature*. 433: 865-8.

Lom J and Schubert G (1983) Ultrastructural study of *Piscinodinium pillulare* (Schäperclaus, 1954) Lom, 1981 with special emphasis on its attachment to the fish host. *J Fish Dis*. 6: 411-428.

López-Garcia P, Philippe H, Gall R, and Moreira D (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers a the Mid-Atlantic Ridge. *Proc Natl Acad Sci.* 100: 697-702.

López-Garcia P, Rodriguez-Valera F, Pedros-Alio C, and Moreira D (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature.* 409: 603-607.

Ludwig M and Gibbs SP (1989) Localization of phycoerythrin at the luminal surface of the thylakoid membrane in *Rhodomonas lens*. *J Cell Biol*. 108: 875-884.

Lynn DH (1988) Cytoterminology of cortical components of ciliates: somatic and oral kinetids. *BioSystems*. 21: 299-307.

Mackin, JG, Malcolm Owen, H, and Collier, A (1950) Preliminary note on the occurance of a new protistan parasite *Dermocystidium marinum* n. sp. in *Crassostrea virginica* (Gmelin). *Science*. 111: 328-329.

Maddison , WP (1997) Gene trees in species trees. *Syst Biol*. 46: 523-536.

Manier, JF, Fize, A, Grizel, H (1971) *Syndium gammari* n. sp. Péridinien Duboscquodinida Syndinidae, parasite de *Gammarus locusta* (Lin.) Crustacé Amphipode. *Protistologica*. 7: 213-219.

Manton I (1967a) Further observations on the fine structure of *Chrysochromulina chiton* with special reference to the haptonema, 'peculiar' Golgi structure and scale production. *J Cell Sci*. 2: 265-272.

Manton I (1967b) Further observations on scale formation in *Chrysochromulina chiton*. *J Cell Sci*. 2: 411-418.

Manton I, Kowallik K, and von Stosch HA (1970) Observations on the fine structure and development of the spindle at mitosis and meiosis in a marine centric diatom (*Lithodesmium undulatum*) IV. The second meiotic division and conclusion. *J Cell Sci*. 7: 407-443.

Maranda, L (2001) Infection of *Prorocentrum minimum* (Dinophyceae) by the parasite *Amoebophrya* sp (Dinoflagellea). *J Phycol*. 37: 245-248.

Margulis L (1996) Archaea-eubacterial mergers in the origin of Eukarya: phylogenetic classification of life. *Proc Natl Accad Sci USA*. 93: 1071-1076.

Marsh ME (2003) Regulation of CaCO3 formation in coccolithophores. *Comp Biochem Phys B*. 136: 743-754.

Martin W and Russell MJ (2003) On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic

prokaryotes, and from prokaryotes to nucleated cells. *Phil. Trans. R. Soc. Lond.* B. 358: 59-85.

Martin W, Stoebe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV (1999) Gene transfer to the nucleus and the evolution of chloroplasts. *Nature*. 393: 162-165.

Mather MW and Vaidya AB (2008) Mitochondria in malaria and related parasites: ancient, diverse and streamlined. *J Bioenerg Biomembr.* 40: 425-433.

Matsuzaki M, Misumi O, Shin-I T, Maruyama S, Takahara M, Miyagishima SY, Mori T, Nishida K, Yagisawa F, Nishida K, Yoshida Y, Nishimura Y, Nakao S, Kobayashi T, Momoyama Y, Higashiyama T, Minoda A, Sano M, Nomoto H, Oishi K, Hayashi H, Ohta F, Nishizaka S, Haga S, Miura S, Morishita T, Kabeya Y, Terasawa K, Suzuki Y, Ishii Y, Asakawa S, Takano H, Ohta N, Kuroiwa H, Tanaka K, Shimizu N, Sugano S, Sato N, Nozaki H, Ogasawara N, Kohara Y, Kuroiwa T (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon* merolae 10D. *Nature.* 428: 653-7.

McFadden GI (1990) Evidence that cryptomonad chloroplasts evolved from photosynthetic eukaryotic endosymbionts. *J Cell Sci*. 95: 303-308.

Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Maréchal-Drouard L, Marshall WF, Qu LH, Nelson DR, Sanderfoot AA, Spalding MH, Kapitonov VV, Ren Q, Ferris P, Lindquist E, Shapiro H, Lucas SM, Grimwood J, Schmutz J, Cardol P, Cerutti H,

Chanfreau G, Chen CL, Cognat V, Croft MT, Dent R, Dutcher S, Fernández E, Fukuzawa H, González-Ballester D, González-Halphen D, Hallmann A, Hanikenne M, Hippler M, Inwood W, Jabbari K, Kalanon M, Kuras R, Lefebvre PA, Lemaire SD, Lobanov AV, Lohr M, Manuell A, Meier I, Mets L, Mittag M, Mittelmeier T, Moroney JV, Moseley J, Napoli C, Nedelcu AM, Niyogi K, Novoselov SV, Paulsen IT, Pazour G, Purton S, Ral JP, Riaño-Pachón DM, Riekhof W, Rymarquis L, Schroda M, Stern D, Umen J, Willows R, Wilson N, Zimmer SL, Allmer J, Balk J, Bisova K, Chen CJ, Elias M, Gendler K, Hauser C, Lamb MR, Ledford H, Long JC, Minagawa J, Page MD, Pan J, Pootakham W, Roje S, Rose A, Stahlberg E, Terauchi AM, Yang P, Ball S, Bowler C, Dieckmann CL, Gladyshev VN, Green P, Jorgensen R, Mayfield S, Mueller-Roeber B, Rajamani S, Sayre RT, Brokstein P, Dubchak I, Goodstein D, Hornick L, Huang YW, Jhaveri J, Luo Y, Martínez D, Ngau WC, Otillar B, Poliakov A, Porter A, Szajkowski L, Werner G, Zhou K, Grigoriev IV, Rokhsar DS, Grossman AR (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*. 318:245-50.

Miller, Lh, Aikawa, M, Johnson, JG, and Shiroishi, T (1979) Interaction between cytochalasin b treated malarial parasites and erythrocytes - attachment and junction formation. *J Exp Med* 149: 172-184.

Mitchell, SF, Ball, JD, Crowley, SF, Marshall, JD, Paul, CRC, Veltkamp, CJ, and Samir, A (1997) Isotope data from Cretaceous chalks and foraminifera: Environmental or diagenetic signals? *Geology*. 25: 691-694.

Montagnes, DJS, Chambouvet, A, Guillou, L, and Fenton, A (2008) Responsibility of microzooplankton and parasite pressure for the demise of toxic dinoflagellate blooms. *Aquat Microb Ecol*. 53: 211-225.

Moon-van der Staay SY, De Wachter R, and Vaulot D (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature*. 409: 607-610.

Moore RB, Obornik M, Janouskovec J, Chrudimsky T, Vancova M, Green DH, Wright SW, Davies NW, Bolch CJS, Heimann K, Slapeta J, Hoegh-Guldberg O, Logsdon JM, Carter DA (2008) A photosynthetic alveolate closely related to apicomplexan parasites. *Nature*. 451: 959-963.

Morrison HG, McArthur AG, Gillin FD, Aley SB, Adam RD, Olsen GJ, Best AA, Cande WZ, Chen F, Cipriano MJ, Davids BJ, Dawson SC, Elmendorf HG, Hehl AB, Holder ME, Huse SM, Kim UU, Lasek-Nesselquist E, Manning G, Nigam A, Nixon JE, Palm D, Passamaneck NE, Prabhu A, Reich CI, Reiner DS, Samuelson J, Svard SG, Sogin ML (2007) Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. *Science.* 317: 1921-6.

Moustafa, A, Beszteri, B, Maier, UG, Bowler, C, Valentin, K, Bhattacharya, D (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science*. 324: 1724-1726

Moustafa, A, Bexzteri, B, Maier, UG, Bowler, C, Valentin, K, Bhattacharya, D (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science*. 1724-1726.

Myl'nikov, AP (2000) The new marine carnivorous flagellate *Colpodella pontica* (Colpodellida, Protozoa). *Zool Zh.* 79: 261-266.

Muscatine L and Porter JW (1997) Reef corals: mutualistic symbioses adapted to nutrient-poor environments. *Biosciences*. 27: 454-460.

Nassoury N, Cappadocia M, and Morse D (2003) Plastid ultrastructure defines the protein import pathway in dinoflagellates. *J Cell Sci*. 116: 2867-2874.

Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol*. 48: 443-53.

Ngo HM, Yang M, and Joiner KA (2004) Are rhoptries in Apicomplexan parasites secretory granules or secretory lysosomal granules? *Mol Microbiol*. 52: 1531-1541.

Nierman WC, Pain A, Anderson MJ, Wortman JR, Kim HS, Arroyo J, Berriman M, Abe K, Archer DB, Bermejo C, Bennett J, Bowyer P, Chen D, Collins M, Coulsen R, Davies R, Dyer PS, Farman M, Fedorova N, Fedorova N, Feldblyum TV, Fischer R, Fosker N, Fraser A, García JL, García MJ, Goble A, Goldman GH, Gomi K, Griffith-Jones S, Gwilliam R, Haas B, Haas H, Harris D, Horiuchi H, Huang J, Humphray S, Jiménez J, Keller N, Khouri H, Kitamoto K, Kobayashi T,

Konzack S, Kulkarni R, Kumagai T, Lafon A, Latgé JP, Li W, Lord A, Lu C, Majoros WH, May GS, Miller BL, Mohamoud Y, Molina M, Monod M, Mouyna I, Mulligan S, Murphy L, O'Neil S, Paulsen I, Peñalva MA, Pertea M, Price C, Pritchard BL, Quail MA, Rabbinowitsch E, Rawlins N, Rajandream MA, Reichard U, Renauld H, Robson GD, Rodriguez de Córdoba S, Rodríguez-Peña JM, Ronning CM, Rutter S, Salzberg SL, Sanchez M, Sánchez-Ferrero JC, Saunders D, Seeger K, Squares R, Squares S, Takeuchi M, Tekaia F, Turner G, Vazquez de Aldana CR, Weidman J, White O, Woodward J, Yu JH, Fraser C, Galagan JE, Asai K, Machida M, Hall N, Barrell B, and Denning DW (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus. Nature. 438:1151-6

Nikoh N and Nakabachi A (2009) Aphids acquired symbiotic genes via lateral gene transfer. *BMC Biology*. **7**:12 doi:10.1186/1741-7007-7-12.

Nishi M, Hu K, Murray JM, and Roos DS (2008) Organellar dynamics during the cell cycle of *Toxoplasma gondii*. *J Cell Sci*. 121: 1559-1568.

Noren, F, Moestrup, O, Rehnstam-Holm, AS (1999) *Parvilucifera infectans* Noren et Moestrup gen. et sp nov (Perkinsozoa phylum nov.): a parasitic flagellate capable of killing toxic microalgae. *Eur J Protistol.* 35: 233-254.

Nosenko T, Lidie KL, Van Dolah FM, Erika Lindquist E, Cheng J-F, US Department of Energy–Joint Genome Institute, and Bhattacharya D (2006)

Chimeric plastid proteome in florida red tide dinoflagellate *Karenia brevis*. *Mol Biol Evol*. 23: 2026-2038.

Not R, Gausling R, Azam F, Heidelberg JF, and Worden AZ (2007) Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ Microbiol*. 9: 1233-1252.

Oborník M, Janouskovec J, Chrudimsky T and Lukes J (2009) Evolution of the apicoplast and its hosts: From heterotrophy to autotrophy and back again. *Int J Parasitol*. 39: 1-12.

O'Kelly CJ (2002) *Glossomastix chrysoplasta* n. gen., n. sp. (Pinguiophyceae), a new coccoidal, colony-forming golden algae from southern Australia. *Phycol Res*. 50: 67-74.

Palmer JD and Delwiche CF (1996) Second-hand chloroplasts and the case of the disappearing nucleus. *Proc Natl Acad Sci.* 93: 7432-7435.

Pascher A (1914) Über flagellaten un algen. Berichter der Deutschen botanischen Gesellschaft 32: 136-160.

Patron NJ, Inagaki Y, and Keeling PJ (2007) Multiple gene phylogenies support the monohyly of cryptomonad and haptophyte host lineages. *Current Biology*. 17: 887-891.

Patron NJ, Waller RF, Archibald JM, and Keeling PJ (2005) Complex protein targeting to dinoflagellates plastids. *J Mol Biol*. 348: 1015-1024.

Patron NJ, Waller RF, and Keeling PJ (2006) A tertiary plastid uses genes from two endosymbionts. *J Mol Biol*. 357: 1373-1382.

Patterson DJ (1999) The diversity of eukaryotes. *Am Nat.* 154: S96-S124.

Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, Peters N, Adlem E, Tivey A, Aslett M, Kerhornou A, Ivens A, Fraser A, Rajandream MA, Carver T, Norbertczak H, Chillingworth T, Hance Z, Jagels K, Moule S, Ormond D, Rutter S, Squares R, Whitehead S, Rabbinowitsch E, Arrowsmith C, White B, Thurston S, Bringaud F, Baldauf SL, Faulconbridge A, Jeffares D, Depledge DP, Oyola SO, Hilley JD, Brito LO, Tosi LR, Barrell B, Cruz AK, Mottram JC, Smith DF, Berriman M (2007) Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat Genet*. 39(7):839-47.

Pearson GA, Hoarau G, Lago-Leston A, Coyer JA, Kube M, Reinhardt R, Henckel K, Serrão ET, Corre E, Olsen JL (2010) An expressed sequence tag analysis of the intertidal brown seaweeds *Fucus serratus* (L.) and *F. vesiculosus* (L.) (Heterokontophyta, Phaeophyceae) in response to abiotic stressors. *Mar Biotechnol* (NY). 12: 195-213.

Pedersen BH, Buchmann K, and Køie, M (1993) Identifying the lethal fish egg parasite *Ichthyodinium chabelardi* as a member of the marine alveolate group I. *Environ Microbiol*. 16: 29-33.

Perkins FO (1996) The structure of *Perkinsus marinus* (Mackin, Owen and Collier, 1950) Levine, 1978 with comments on taxonomy and phylogeny of *Perkinsus* spp. *J Shellfish Res.* 15:67-87.

Perkins FO. (1976). Zoospores of oyster pathogen, *Dermocystidium marinum*. I. Fine structure of conoid and other sporozoan-like organelles. *J Parasitol.* 62: 959-974.

Piaggio-Talice R, Burleigh G, and Eulenstein O (2004) Quartet Supertrees, *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, pages 173--192. Springer.

Preiser P, Kaviratne M, Khan S, Bannister L, and Jarra W (2000) The apical organelles of malaria merozoites: host cell selection, invasion, host immunity and immune evasion. *Microbes and Infection*. 2: 1461-1477.

Reichle RE (1969) Fine structure of *Phytophthora parasitica* zoospores. *Mycologia*. 61: 30-51.

Reith M (1995) Molecular-biology of rhodophyte and chromophyte plastids. *Annu Rev Plant Physiol Plant Mol Biol*. 46: 549-575.

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y, Tanahashi T, Sakakibara K, Fujita T, Oishi K, Shin-I T, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S, Yamaguchi K, Sugano S, Kohara Y, Fujiyama A, Anterola A, Aoki S, Ashton N, Barbazuk WB,

Barker E, Bennetzen JL, Blankenship R, Cho SH, Dutcher SK, Estelle M, Fawcett JA, Gundlach H, Hanada K, Heyl A, Hicks KA, Hughes J, Lohr M, Mayer K, Melkozernov A, Murata T, Nelson DR, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton PJ, Sanderfoot A, Schween G, Shiu SH, Stueber K, Theodoulou FL, Tu H, Van de Peer Y, Verrier PJ, Waters E, Wood A, Yang L, Cove D, Cuming AC, Hasebe M, Lucas S, Mishler BD, Reski R, Grigoriev IV, Quatrano RS, Boore JL (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science*. 319: 64-9.

Rice P, Longden I, and Bleasby A (2000) EMBOSS: The European molecular biology open software suite. *Trends in Genetics*. 16: 276-277.

Riebesell U, Zondervan I, Rost B, Tortell PD, Zeebe RE, and Morel FMM (2000) Reduced calcification of marine plankton in response to increased atmospheric CO_2. *Nature* 407: 364-367.

Riisberg I, Orr RJ, Kluge R, Shalchian-Tabrizi K, Bowers HA, Patil V, Edvardsen B, Jakobsen KS. (2009) Seven gene phylogeny of heterokonts. *Protist*. 160: 191-204.

Ris H, and Kubai DF (1974) Unusual mitotic mechanism in parasitic protozoan *Syndinium* sp. *J Cell Biol*. 60(3): 702-720.

Rivera MC, Jain, R, Moore JE, and Lake, JA (1998) Genomic evidence for two functionally distinct gene classes. *Proc Natl Acad Sci*. 95: 6239-6244.

Rizzo, PJ (2003) Those amazing dinoflagellate chromosomes. *Cell Res*. 13: 215-217.

Rodríguez-Ezpeleta N, Brinkmann H, Burger G, Roger AJ, Gray MW, Philippe H, Lang BF (2007) Toward resolving the eukaryotic tree: the phylogenetic positions of jakobids and cercozoans. *Curr Biol*. 17: 1420-5.

Rodriquez-Lanetty M and Hoegh-Guldberg O (2002) The phylogeography and connectivity of the latitudinally widespread scleractinian coral *Plesiastrea versipora* in the Western Pacific. Mol Ecol. 11: 1177-1189.

Robinson ND (1995) Biogenic acid rain during the late Cretaceous as a possible cause of extinctions. *J Geol Soc London*. 152: 4-6.

Rumpho ME, Worful JM, Lee J, Kannan K, Tyler MS, Bhattacharia D, Moustafa A, and Manhart (2008) Horizontal gene transfer of the algal nuclear gene psbO to the photosynthetic sea slug *Elysia chlorotica*. *PNAS*. 105: 17867-17871.

Ryning FW and Remington JS (1978) Effect of cytochalasin-D on *Toxoplasma gondii* cell entry. *Infect Immun.* 20: 739-743.

Saffo MB (1993) Coming to terms with a field – words and concepts in symbiosis. *Symbiosis*. 14: 17-31.

Saldarriaga, JF, McEwan, ML, Fast, NM, Taylor, FJ, Keeling, PJ (2003 Jan) Multiple protein phylogenies show that *Oxyrrhis marina* and *Perkinsus marinus*

are early branches of the dinoflagellate lineage. *Int J Syst Evol Microbiol.* 53: 355-65.

Santore UJ (1985) A cytological survey of the genus *Cryptomonas* (Cryptophyceae) with comments on its taxonomy. *Archiv für Protistenkunde*. 130: 1-52.

Schnepf E and Elbrachter M (1992) Nutritional strategies in dinoflagellates. *Eur J Protistol*. 28: 3-24.

Schnepf E and Elbrachter M (1999) Dinophyte chloroplasts and phylogeny – A review. *Grana*. 38: 81-97.

Schnepf, E, Deichgräber, G, Drebes, G (1985) Food uptake and fine structure of the dinophyte *Paulsenella* sp., an ectoparasite of marine diatoms. *Protoplasma.* 124: 188-204.

Sea Urchin Genome Sequencing Consortium (2006) The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science*. 314: 941-52.

Seeber F, Limenitakis J, Soldati-Favre (2008) Apicomplexan mitochondrial metabolism: a story of gains, losses, and retentions. *Trends Parasitol*. 24:468-478.

Shalchlan-Tabrizi K, Brate J, Logares R, Klaveness D, Berney C, and Jakobsen, KS (2008) Diversification of unicellular eukaryotes: cryptomonad colonizations of

marine and fresh waters inferred from revised 18S rRNA phylogeny. *Environ Microbiol*. 10:2635-2644.

Shi T. and Falkowski, PG (2008) Genome evolution in cyanobacteria: the stable core and the variable shell. *Proc Natl Acad Sci*. 105: 2510-2515.

Sibley LD (2004) Intracellular parasite invasion strategies. *Science*. 304: 248-253.

Simpson, AGB Patterson DJ (1996) Ultrastructure and identification of the predatory flagellate *Colpodella pugnax* Cienkowski (Apicomplexa) with a description of *Colpodella turpis* n. sp. and a review of the genus. *Systematic Parasitology.* 33: 187-198.

Skovgaard, A, Massana, R, and Saiz, E (2007) Parasitic species of the genus *Blastodinium* (Blastodiniphyceae) are peridinoid dinoflagellates. *J Phycol.* 43: 553-560.

Skovgaard, A, Massana, R, Balague, V, Saiz, E (2005) Phylogenetic position of the copepod-infesting parasite *Syndinium turbo* (Dinoflagellata, Syndinea). *Protist.* 156: 413-423.

Skovgaard, A, Meneses, I, and Angelico, MM (2009) Identifying the lethal fish egg parasite *Ichthyodinium chabelardi* as a member of Marine Alveolate Group I. *Environ Microbiol.* 11: 2030-2041.

Smith TF and Waterman MS (1981) Identification of common molecular subsequences. *J Mol Biol*. 147: 195-197.

Soyer, MO (1970) Les ultrastructures liées aux fonctions de relation chez *Noctiluca miliaris* S. (Dinoflagellata). *Zeitschrift für Zellforschung und Mikroskopische Anatomie* 104: 29-55.

Soyer, MO (1972) Les ultrastructures nucléaires de la Noctiluque (Dinoflagellé libre) au cours de la sporogenése. *Chromosoma*. 39: 419-441.

Soyer, MO (1974) Étude ultrastructurale de *Syndinium* sp. Chatton parasite coelomique de copépodes Pélagiques. *Vie Milieu.* 24: 191-212.

Stajich J, Block D, Boulez K, Brenner S, Chervitz S, Dagdigian C, Fuellen G, Gilbert J, Korf I, Lapp H, Lehväslaiho H, Matsalla C, Mungall C, Osborne B, Pocock M, Schattner P, Senger M, Stein L, Stupka E, Wilkinson M, Birney E (2002). "The Bioperl toolkit: Perl modules for the life sciences". *Genome Res* 12 (10): 1611-1618.

Stamatakis, A (2006) RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic Analyses with Thousands of Taxa and Mixed Models. *Bioinformatics.* 22: 2688–2690.

Stentiford, GD, Shields, JD (2005) A review of the parasitic dinoflagellates *Hematodinium* species and *Hematodinium*-like infections in marine crustaceans. *Dis Aquat Organ.* 66: 47-70.

Strain HH, Manning WM, Hardin G (1944) Xanthophylls and carotenes of diatoms, brown algae, dinoflagellates, and sea anemones. *Biol Bull*. 86: 169-171.

Struck NS, de Souza Dias S, Langer C, Marti M, Pearce JA, Cowman AF, Gilberger TW (2005) Re-defining the Golgi complex in *Plasmodium falciparum* using the novel Golgi marker PfGRASP. *J Cell Sci*. 118: 5603-13.

Sunila, I, Hamilton, RM, Dungan, CF (2001) Ultrastructural characteristics of the In Vitro cell cycle of the protozoan pathogen of oysters, *Perkinsus marinus*. *J Eukaryot Microbiol.* 48: 348-361.

Sweeny, BM (1978) Ultrastructure of *Noctiluca miliaris* (Pyrrophyta) with green flagellate symbionts *J Phycol.* 14: 116-120.

Tatusov, RL, Koonin, EV, Lipman, DJ (1997) A Genomic Perspective on Protein Families. *Science* 278, 631-637.

Taylor, FJR (1968) Parasitism of toxin-producing dinoflagellate *Gonyaulax catenella* by endoparasitic dinoflagellate *Amoebophrya ceratii*. *Journal Fish Research Bd Canada.* 25: 2241-2245.

Taylor FJR (1980) On dinoflagellate evolution. *BioSystems*. 13: 65-108;

Tengs T, Dahlberg OJ, Shalchian-Tabrizi K, Klaveness D, Rudi K, Delwiche CF, and Jakobsen KS (2000) Phylogenetic analyses indicate that the

19'Hexanoyloxy-fucoxanthin-containing dinoflagellates have tertiary plastids of haptophyte origin. *Mol Biol Evol*. 17: 718-729.

Thrall PH, Hochberg ME, Burdon JJ, and Bever JD (2006) Coevolution of symbiotic mutualists and parasites in a community context. *Trends Ecol Evol*. 22: 120-126.

Tillmann U (1998) Phagotrophy by a plastidic haptophyte, *Prymnesium patelliferum*. *Aquat Microb Ecol*. 14: 155-160.

Tyler BM, Tripathy S, Zhang X, Dehal P, Jiang RH, Aerts A, Arredondo FD, Baxter L, Bensasson D, Beynon JL, Chapman J, Damasceno CM, Dorrance AE, Dou D, Dickerman AW, Dubchak IL, Garbelotto M, Gijzen M, Gordon SG, Govers F, Grunwald NJ, Huang W, Ivors KL, Jones RW, Kamoun S, Krampis K, Lamour KH, Lee MK, McDonald WH, Medina M, Meijer HJ, Nordberg EK, Maclean DJ, Ospina-Giraldo MD, Morris PF, Phuntumart V, Putnam NH, Rash S, Rose JK, Sakihama Y, Salamov AA, Savidor A, Scheuring CF, Smith BM, Sobral BW, Terry A, Torto-Alalibo TA, Win J, Xu Z, Zhang H, Grigoriev IV, Rokhsar DS, Boore JL (2006) *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science*. 313: 1261-6

Ucko, M, Elbrächter, M, Schnepf, E (1997) A *Crypthecodinium cohnii*-like dinoflagellate feeding myzocytotically on the unicellular red alga *Porphyridium* sp. *Eur J Phyco*. 32: 133-140.

Urushihara, H (2008) Developmental biology of the social amoeba: history, current knowledge and prospects. *Develop Growth Differ*. 50: S277-S281.

Van der Wal, P, Kempers, RS, and Veldhuis, MJW (1995) Production and downward flux of organic matter and calcite in a North Sea bloom of the coccolithophore *Emiliania huxleyi*. *Mar Ecol Prog Ser* 126: 247-265.

Van Dooren GG, Stimmler LM, and McFadden, GI (2006) Metabolic maps and functions of the *Plasmodium* mitochondrion. *FEMS Microbiol Rev.* 30: 596-630.

Venable JH and Coggeshall R (1965) A simplified lead citrate stain for use in electron microscopy. *J Cell Biol.* 25: 407-408.

Villalba A, Reece KS, Ordas MC, Casas SM, and Figueras A (2004) Perkinsosis in molluscs: A review. *Aquat Living Resour*. 17: 411-432.

Waller RF, Keeling PJ, Donald RGK, Striepen B, Handman E, Lang-Unnasch N, Cowman AF, Besra GS, Roos DS and McFadden GI (1998) Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. *Proc Natl Acad Sci.* 95: 12352-12357.

Wang D (2008) Neurotoxins from marine dinoflagellates: a brief review. *Mar Drugs*. 6: 349-371.

Wellner A, Lurie MN, and Gophna U (2007) Complexity, connectivity, and duplicatability as barriers to lateral gene transfer. *Genome Biology* 2007, 8:R156 (doi:10.1186/gb-2007-8-8-r156).

Wetzel DM, Schmidt J, Kuhlenschmidt MS, Dubey JP, and Sibley LD (2005) Gliding motility leads to active cellular invasion by *Cryptosporidium parvum* sporozoites. *Infect immune*. 73: 5379-5387.

Whatley JM, John P, and Whatley FR (1979) From extracellular to intracellular: the establishment of mitochondria and chloroplasts. *Proc R Soc Lond B*. 204: 165-187.

Whelan S and Goldman N (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*. 18: 691-699.

Whitaker JW, McConkey GA, and Westhead DR (2009) The tranferome of metabolic genes explored: analysis of the horizontal transfer of enzyme encoding genes in unicellular eukaryotes. *Genome Biology*. :(doi:10.1186/gb-2009-10-4-r36).

Wilcox LW and Wedemayer GJ (1991) Phagotrophy in fresh-water, photosynthetic dinoflagellate *Amphidinium cryophilum*. *J. Phycol*. 27: 600-609.

Wolters J (1991) The troublesome parasites--molecular and morphological evidence that Apicomplexa belong to the dinoflagellate-ciliate clade. *Biosystems.* 25: 75-83.

Wong JTY and Kwok ACM (2005) Proliferation of dinoflagellates: blooming or bleaching. *Bioessays* 27: 730-740.

Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A, Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, Hornsby T, Howarth S, Huckle EJ, Hunt S, Jagels K, James K, Jones L, Jones M, Leather S, McDonald S, McLean J, Mooney P, Moule S, Mungall K, Murphy L, Niblett D, Odell C, Oliver K, O'Neil S, Pearson D, Quail MA, Rabbinowitsch E, Rutherford K, Rutter S, Saunders D, Seeger K, Sharp S, Skelton J, Simmonds M, Squares R, Squares S, Stevens K, Taylor K, Taylor RG, Tivey A, Walsh S, Warren T, Whitehead S, Woodward J, Volckaert G, Aert R, Robben J, Grymonprez B, Weltjens I, Vanstreels E, Rieger M, Schäfer M, Müller-Auer S, Gabel C, Fuchs M, Düsterhöft A, Fritzc C, Holzer E, Moestl D, Hilbert H, Borzym K, Langer I, Beck A, Lehrach H, Reinhardt R, Pohl TM, Eger P, Zimmermann W, Wedler H, Wambutt R, Purnelle B, Goffeau A, Cadieu E, Dréano S, Gloux S, Lelaure V, Mottier S, Galibert F, Aves SJ, Xiang Z, Hunt C, Moore K, Hurst SM, Lucas M, Rochet M, Gaillardin C, Tallada VA, Garzon A, Thode G, Daga RR, Cruzado L, Jimenez J, Sánchez M, del Rey F, Benito J, Domínguez A, Revuelta JL, Moreno S, Armstrong J, Forsburg SL, Cerutti L, Lowe T, McCombie WR, Paulsen I, Potashkin J, Shpakovski GV, Ussery D, Barrell BG, Nurse P (2002) The genome sequence of *Schizosaccharomyces pombe. Nature.* 415: 871-80.

World Health Organization (2008) World malaria report 2008 p vii.

Yellowlees D, Rees TA, Leggat W (2008) Metabolic interactions between algal symbionts and invertebrate hosts. *Plant Cell Environ*. 31: 679-694.

Yoon HS, Hackett JD, Ciniglia C, Pinto G, and Bhattacharya D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol*. 21: 809-18.

Yoon HS, Hackett JD, Pinto G, and Bhattacharya D (2002) The single, ancient origin of chromist plastids. *Proc Natl Acad Sci*. 99: 15507-15512.

Yoshida M, Noël M-H, Nakayama T, Naganuma T, and Inouye I (2006) A haptophyte bearing siliceous scales: Ultrastructure and phylogenetic position of *Hyalolithus neolepis* gen. et sp. nov. (Prymnesiophyceae, Haptophyta). *Protist*. 157: 213-234.